arXiv:2602.04237v1 [math.OC] 4 Feb 2026

用于非光滑函数的改进的增强型 DC 算法及其在图像恢复中的应用

ZeyuLi[1†]、Te Qi[1†]、Tieyong Zeng[2, 3*]

[1*]香港中文大学数学系，香港沙田，999077。[2]北京师范大学-香港浸会大学高等研究院，中国广东省珠海市，519087。[3]广州南方学院数学与统计学院，中国广东省广州市，510970。

通讯作者。电子邮箱：tieyongzeng@bnbu.edu.cn；贡献作者：1155184076@link.cuhk.edu.hk；tqi@math.cuhk.edu.hk；

†这些作者对本研究的贡献相同。

摘要

我们提出一种新方法，用于在涉及凸函数之差（DC）的非光滑非凸问题上执行加速凸差分函数算法（BDCA）。近期提出的BDCA通过线搜索过程，沿下降方向从经典差分凸函数算法（DCA）计算出的点进行外推，以进一步降低目标函数值并加速DCA的收敛。然而，当DC分解中的第一个函数非光滑时，BDCA计算出的方向可能是上升方向，因此无法进行单调线搜索。本文针对某些类型的非光滑DC规划提出了一种单调的改进增强差分凸函数算法（IBDCA），即那些可表示为一个可能非光滑函数与一个光滑函数之差的问题。我们证明，IBDCA生成的序列的任一聚点都是所研究问题的临界点，且相应的目标函数值单调递减并收敛。我们还在Kurdyka–Lojasiewicz性质下给出了全局收敛性和收敛速率。IBDCA在图像恢复中的应用表明了所提方法的有效性。相应的数值实验表明

我们的IBDCA在计算时间和迭代次数两方面均优于DCA及其他先进的DC方法。

关键词： DC规划、加速凸差分函数算法、全局收敛、图像恢复。

1 引言

本文主要关注差凸（DC）优化问题

$$\min_{x\in\mathbb{R}^m} \phi(x) := g(x) - h(x), \tag{1}$$

其中 g，h：$\mathbb{R}^m \to \mathbb{R}\cup\{+\infty\}$ 是适当的凸函数，且 $\phi(x)$满足：

$$\inf_{x\in\mathbb{R}^m} \phi(x) > -\infty. \tag{2}$$

在此，我们还对 ∞ 的定义作如下约定 [1, 2]:

$$\begin{gathered}(+\infty)-(+\infty)=(+\infty)-(+\infty)=+\infty,\\ (+\infty)-c=+\infty \qquad \text{and} \qquad c-(+\infty)=-\infty, \qquad \forall c\in\,]\infty,+\infty[\,.\end{gathered}$$

众所周知的凸函数之差算法（DCA） [3, 4] 是求解（1）的最高效算法之一。对DCA的广泛研究与发展已经持续了30多年 [5–7]。DCA已广泛应用于各个领域，包括机器学习、稀疏优化、深度学习和图像处理 [8–11]。多年来，DCA的若干变体在理论和实践方面都得到了探索。DCA的代表性变体包括次梯度法 [12, 13],近端算法 [14–16], 以及惯性方法 [17]。

近年来， [1, 18] 提出了一种加速DC算法（BDCA）来求解（1），并显著加快了DCA的收敛速度。其中， [18] 关注的是（ g 和 h均为光滑函数的情形。1）。 [1]随后将其推广到 g 为光滑函数而 h 可能非光滑的情形。BDCA基于DCA计算得到的点，沿下降方向额外执行一次线搜索。除了提高DCA的收敛速度外，BDCA还可以借助线搜索步骤有效避免算法陷入局部最优解，从而增加收敛到全局最优解的可能性[1,例3。1]。

一种特殊类型的DC问题经常出现在图像与信号处理领域的各种模型中：其DC分解的第一项可能是非光滑的，而第二项是光滑的。相关应用包括图像去噪 [19–22], 非凸正则化 [23–25], 压缩感知 [26], 相位恢复[27], 以及字典学习 [28]。然而，当函数 g 非光滑时，BDCA中的搜索方向可能成为上升方向（见 [1, 例3.2] 和[18,注3.1]）。因此，BDCA可能无法有效处理这类DC问题。为解决这一问题，Ferreira等人 [29] 提出了一种非单调增强

差分凸函数算法（nmBDCA），并进一步将nmBDCA应用于图像去噪 [30]。该算法采用非单调线搜索，适用于DC函数的两个组成部分均非光滑的情形。nmBDCA允许目标函数值在一定范围内增长，并通过非单调线搜索步骤中的特定参数对增长进行控制。研究人员还针对特定模型开发了其他几种加速DC算法。例如， [31] 通过平方和分解，对投资组合模型中的投影差分凸方法进行增强，该模型涉及前四阶矩。 [32] 基于幂和DC分解，采用精确线搜索加速DCA，以求解线性约束多项式规划。

本文旨在提出一种改进的加速凸差分函数算法（IBDCA），当仅能保证第二个 DC 分量光滑时，通过单调线搜索对 DCA 进行外推。一个关键观察是，在这类 DC问题中，BDCA搜索方向可能使当前迭代点上升，但对上一次 DCA迭代生成的点而言却是下降方向。基于这一观察，我们提出的 IBDCA 的核心思想是：从上一次 DCA迭代生成的点出发，沿下降方向进行线搜索，而不像其他 BDCA 类方法那样从当前迭代点出发。通过设置适当的参数和搜索准则，我们的算法不再像 BDCA 那样，在搜索方向可能导致上升的点处执行额外的搜索步骤，而是采用 DCA 的步长；对于搜索方向为下降方向的点，则使用与 BDCA 类似的线搜索和步长。为展示该方法的有效性，我们将 IBDCA 应用于图像恢复中的一个重要高维 DC问题：柯西噪声去除。数值实验表明，在复原效果和计算时间方面，IBDCA 均优于 DCA 及其他先进算法。此外，与 BDCA 类似，IBDCA 比 DCA 和 nmBDCA 更有可能摆脱较差的局部解。

值得一提的是，BDCA采用了[33, 34], 中较早提出的搜索方向，并且他们也提出了从上一迭代生成的点出发进行线搜索的想法。然而，Fukushima-Mine（F-M）方法与我们提出的方法有很大不同。实际上，F-M方法的最大步长等于1（与DCA相同），而IBDCA在许多情况下可以加速DCA。我们证明，在许多情形下，所提方法的线搜索可以取得与BDCA相似的结果。另一方面，在 [34], 中，Fukushima-Mine方法既没有给出目标函数下降的准确估计，也没有规定具体的线搜索准则。相比之下，我们证明，在所设计的搜索准则下，我们的方法展现出与BDCA相似的收敛性和下降性质。因此，所提出的IBDCA将BDCA推广到更广泛的一类非光滑目标函数。我们还基于Kurdyka–-Lojasiewicz（K-L）性质给出了IBDCA的全局收敛和收敛速率，该性质涵盖了大量应用模型，例如图像处理。

本文其余部分的组织结构如下。在第2节中，我们回顾一些预备概念和结果，并给出全文所采用的假设。所提出的 IBDCA 及其基本性质将在第3节中介绍。其收敛性

分析和 K-Ł 性质下的收敛速率将在第4节中给出。在第5节中，我们将介绍 IBDCA 在图像恢复中的应用，并比较 IBDCA 与最先进的算法的性能。最后，在第6节中，我们将简要讨论一些结论和未来研究方向。

2 预备知识

本节回顾后文将用到的一些预备概念和结果，并陈述我们对 (1) 所作的假设。本文中，两个向量 $x, y \in \mathbb{R}^m$ 的内积记为 $\langle x,y\rangle$，而 $\|\cdot\|$ 表示诱导范数。我们用 $\mathbb{B}(x,r)$ 表示以 x 为中心、半径为 $r>0$的开球。函数 $f\colon \mathbb{R}^m \to \mathbb{R}\cup\{+\infty\}$ 在某点$x \in \mathbb{R}^m$ 可微时，其梯度记为 $\nabla f(x)$。

2.1 凸分析中的基本概念与结果

本小节回顾凸分析以及非光滑函数广义导数的一些基本定义和相关结论。

定义 1 设 $f\colon \mathbb{R}^m \to \mathbb{R}\cup\{+\infty\}$ 是一个扩展实值函数，其中 $\mathbb{R}\cup\{+\infty\}$ 表示扩展实数线。f 的定义域定义为

$$\operatorname{dom} f = \left\{x \in \mathbb{R}^m : f(x) < +\infty\right\}.$$

f 称为正常的，如果$\operatorname{dom} f \neq \emptyset$。

定义 2 称函数 $f\colon \mathbb{R}^m \to \mathbb{R}\cup\{+\infty\}$ 为凸的，如果

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y) \quad \text{for all } x, y \in \mathbb{R}^m \text{ and } \lambda \in [0,1].$$

f 称为以模数 $\rho > 0$ 强凸的，如果$f - \frac{\rho}{2}\|\cdot\|^2$ 是凸的。 _

定义 3 称函数 $F\colon \mathbb{R}^m \to \mathbb{R}^n$ 为利普希茨连续的，如果存在某个常数 $L \ge 0$ 使得

$$\|F(x) - F(y)\| \le L\|x - y\|, \quad \text{for all } x, y \in \mathbb{R}^m.$$

F 称为局部利普希茨的，如果对每一个 $x \in \mathbb{R}^m$，存在 x 的一个邻域 U_x，使得 F 限制在 U_x 上是利普希茨连续的。

命题 1 ([35]) *Let $f\colon \mathbb{R}^m \to \mathbb{R}\cup\{+\infty\}$ be an extended real-valued function, then f is convex if and only if*

$$\langle u - v, x - y\rangle \ge 0, \quad \text{for all } x, y \in \mathbb{R}^m \text{ and } u \in \partial f(x), v \in \partial f(y).$$

f is ρ-strongly convex ($\rho > 0$) if and only if

$$\langle u - v, x - y\rangle \ge \rho\|x - y\|^2, \quad \text{for all } x, y \in \mathbb{R}^m \text{ and } u \in \partial f(x), v \in \partial f(y).$$

定义 4 设 $f:\ \mathbb{R}^m \to \mathbb{R}\cup\{+\infty\}$ 为局部利普希茨连续，则 f 在 $\tilde{x}$ 处的克拉克次微分 $\partial_C f(\tilde{x})$ 定义为

$$\partial_C f(\tilde{x}) = \mathrm{co}\left\{\lim_{x\to\tilde{x}, x\in D_f} \nabla f(x)\right\},$$

其中，co 表示凸包，D_f 是 f 的不可微点集合。根据拉德马赫定理，该集合的勒贝格测度为零 [36]。

命题 2（克拉克次微分的性质，[37]） *(1) If f is convex in a neighborhoodof x, then*

$$\partial_C f(x) = \partial f(x).$$

(2) If f is differentiable at x, then we have

$$\partial_C f(x) = \{\nabla f(x)\}.$$

(3) For any constant $s\in\mathbb{R}$, we have

$$\partial_C (sf)(x) = s\partial_C f(x).$$

(4)

$$\partial_C (f+g)(x) \subseteq \partial_C f(x) + \partial_C g(x).$$

In addition, the equality holds if either f or g is differentiable at x.

定义 5 设 $f\colon \mathbb{R}^m \to \mathbb{R}\cup\{+\infty\}$ 是一个扩展实值函数。对于 $x\in \mathrm{dom}\, f$ 和 $d\in\mathbb{R}^m$，f 在 x 沿方向 d 的方向导数定义为:

$$f'(x;d) := \lim_{t\downarrow 0} \frac{f(x+td)-f(x)}{t}.$$

2.2 假设

本文始终对 g 和 h 在（1）中施加以下三个假设条件。

假设1 g 和 h 均为模数为 $\rho>0$的强凸函数;

假设2 对于任意 $x\in\mathrm{dom}\, g$ 和任意 $d\in\mathbb{R}^m$,

$$g'(x;d) = \sup_{v\in\partial g(x)} \langle v, d\rangle; \tag{3}$$

假设3 函数 h 在包含$\mathrm{dom}\, g$ 的某个开集上连续可微，且具有局部利普希茨梯度，并且

$$\inf_{x\in\mathbb{R}^m} \phi(x) > -\infty. \tag{4}$$

Remark 1 假设1 并非限制性条件，因为任意一个 ϕ 的 DC 分解 $\phi = g-h$ 都可以表示为 $\phi = \left(g+\frac{\rho}{2}\|\cdot\|^2\right) - \left(h+\frac{\rho}{2}\|\cdot\|^2\right)$。等式（3）对所有 $x\in\mathrm{ridom}\, g$ 均成立（参见 [38]）。特别地，若 $\mathrm{dom}\, g = \mathbb{R}^m$，则假设2 显然成立。（4）在 DC 规划中很常见，例如参见 [1, 14, 18]。

在这些假设条件下，以下必要最优性条件成立。

命题 3 [33] *If $x^* \in$ ϕ is a local* （一阶必要最优性条件，） dom (), ,

minimizer of eq. 1 then x^* is a critical point i.e.

$$\nabla h(x^*) \in \partial g(x^*).$$

3 单调增强 DC 算法

本节介绍用于求解方程（1）的 IBDCA，并进一步给出该算法的收敛性。

3.1 改进的BDCA

我们首先介绍 DCA 和 BDCA 的概念性表述。DCA 的主要思想是利用问题方程（ h 1）中 的仿射主化来近似f（x），然后对所得凸函数进行最小化。DCA 的主要步骤见算法 1。

算法 1 DCA（（DC算法， [3])

1: 设 x^0 为任意初始点，选择任意 $y^0 \neq x^0$ 并设置 $k :=0$。2: 求解强凸最小化问题，得到唯一解 y^k:

$$\min_{x\in\mathbb{R}^m} g(x) - \left\langle \nabla h\left(x^k\right), x\right\rangle$$

3: 如果 $y^k = x^k$，则停止并返回 x^k。否则，设置 $x^{(k+1)} :=y^k$，设置 $k :=k+1$，然后返回上一步。

BDCA的主要思想是沿着每次DCA迭代所获得的方向 d^k 进行线搜索，以扩展DCA。BDCA的主要步骤如算法2所示。

[1] 指出，如果DC分解中的第一个函数是非光滑的，则BDCA中的线搜索可能不适用。nmBDCA的思想是允许线搜索采用非单调方式，在此过程中目标函数值可能增加。nmBDCA的主要步骤如算法 3所示。然而，当 d^k 的方向为上升方向时，nmBDCA的线搜索无法降低目标函数值。因此，我们提出一种单调递减的改进的BDCA来求解此类DC问题。

我们的 IBDCA 使用与 BDCA 相同的线搜索方向。然而，主要区别在于，B DCA 的线搜索从 y^k开始，而 IBDCA 将 x^k选作线搜索的起点。当 $\lambda^k = 1$ 出现在算法 4的第 4 步时，此次迭代等同于执行一步 DCA，而不执行 BDCA 额外的下降步。这使我们能够避免搜索方向上升时无法应用 BDCA 的情况。此外，我们稍后将指出，我们的

算法 2 BDCA（加速DC算法，[18]）

1: 固定 $\alpha > 0$ 和 $0 < \beta < 1$。令 x^0 为任意初始点，并设 $k := 0$。2: 求解该强凸最小化问题，得到唯一解 y^k:

$$\min_{x \in \mathbb{R}^m} g(x) - \left\langle \nabla h\left(x^k\right), x \right\rangle$$

3: 令 $d^k := y^k - x^k$。若 $d^k = 0$，则停止并返回 x^k。否则，进入下一步。4: 取 $\lambda_k > 0$ 并令 $\lambda_k := \lambda_{k0}$ 当 $\phi(y^k + \lambda_k d^k) > \phi(y^k) - \alpha\lambda_k \left\|d^k\right\|^2$ 时，令 $\lambda_k := \beta\lambda_{k0}$。5: 令 $x^{k+1} := y^k + \lambda_k d^k$。若 $x^{k+1} := x^k$，则停止并返回 x^k；否则，令 $k := k+1$ 并转到步骤 2。

算法 3 nmBDCA（非单调增强型 DC 算法，[39]）

1: 固定 $\alpha > 0$ 和 $0 < \beta < 1$。令 x^0 为任意初始点，并令 $k := 0$。2: 求解强凸最小化问题，以得到唯一解 y^k:

$$\min_{x \in \mathbb{R}^m} g(x) - \left\langle \nabla h\left(x^k\right), x \right\rangle$$

3: 令 $d^k := y^k - x^k$。若 $d^k = 0$，则停止并返回 x^k。否则，进入下一步。4: 取 $\lambda_k > 0$，令 $\lambda_k := \lambda_k$，并选择 $v_k > 0$。当 $\phi(y^k + \lambda_k d^k) > \phi(y^k) - \alpha\lambda_k \left\|d^k\right\|^2 + v_k$ 时，令 $\lambda_k := \beta\lambda_{k0}$。5: 令 $x^{k+1} := y^k + \lambda_k d^k$。若 $x^{k+1} := x^k$，则停止并返回 x^k；否则，令 $k := k+1$ 并转到步骤 2。

算法 4 IBDCA（改进的增强型 DC 算法）

1: 固定 $\bar{\lambda} > 1$ 和 $0 < \beta < 1$。令 x^0 为任意初始点，并令 $k := 0$。2: 求解强凸最小化问题，以得到唯一解 y^k:

$$\min_{x \in \mathbb{R}^m} g(x) - \left\langle \nabla h\left(x^k\right), x \right\rangle \tag{5}$$

3: 设置 $d^k := y^k - x^k$。如果 $d^k = 0$，则停止并返回 x^k。否则，进入下一步。4: 取 $\lambda_k > 0$ 并选择 $\alpha_k > 0$。处理 $\lambda_k := \beta\lambda_k$，直到 λ^k 满足:

$$\phi(x^k + \lambda_k d^k) \leq \phi(x^k) - \alpha\lambda_k \left\|d^k\right\|^2, \tag{6}$$

and

$$\phi(x^k + \lambda_k d^k) \leq \phi(y^k). \tag{7}$$

设置 $\lambda^k = 1$，如果 $\lambda^k \leq 1$，则进入下一步。

5: 令 $x^{k+1} := x^k + \lambda_k d^k$。若 $x^{k+1} := x^k$，则停止并返回 x^k；否则，令 $k := k+1$ 并转到步骤 2。

所提出的 IBDCA 是 BDCA 的一种推广，在许多情况下，它可以像 BDCA 一样执行一个额外的下降步。

3.2 IBDCA 的基本性质

我们的 IBDCA 与 DCA 类似，具有以下性质：

命题 4 *For any $k \in \mathbb{N}$, the following properties hold:*

1. *If $d^k = 0$, then x^k is a critical point of ϕ.*
2. *$\phi(y^k) \leq \phi(x^k) - \rho \left\| d^k \right\|^2$ holds.*

Proof 对于第一个断言，由于 y^k 是方程（5）中强凸问题的唯一解，因此有：

$$\nabla h(x^k) \in \partial g(y^k). \tag{8}$$

If $d^k = y^k - x^k = 0$, then

$$0 \in \partial g(x^k) - \nabla h(y^k) = \partial \phi(x^k).$$

对于第二个断言，利用 ρ 对 g 和 h 的强凸性，我们得到：

$$g(x^k) - g(y^k) \geq -\langle v, d^k \rangle + \rho/2 \left\| d^k \right\|^2, \forall v \in \partial g(y^k),$$

$$h(y^k) - h(x^k) \geq \langle \nabla h(x^k), d^k \rangle + \rho/2 \left\| d^k \right\|^2.$$

然后，对这两个不等式求和并应用式（8），即可得出结论。 □

下面的结果表明，d^k 是 x^k处 ϕ 的下降方向，从而为 IBDCA 提供了一种合理的线搜索策略。

命题 5 *For any $k \in \mathcal{N}$, the following properties hold:*

1. $\phi' \left(x^k;\ d \right) \leq -\rho \left\| d^k \right\|^2$;

2. $\forall\, 0 \leq \alpha < \rho,\ \exists\, \sigma_k > 0,$ *eq.* (6) *holds for any* $\lambda_k \in]0,\ \sigma_k]$.

Proof 根据单侧方向导数的定义和假设 2，我们有：

$$\begin{aligned}
\phi' \left(x^k; d^k \right) &= \lim_{t \downarrow 0} \frac{\phi \left(x^k + t d^k \right) - \phi \left(x^k \right)}{t} \\
&= \lim_{t \downarrow 0} \frac{g \left(x^k + t d^k \right) - g \left(x^k \right)}{t} - \lim_{t \downarrow 0} \frac{h \left(x^k + t d^k \right) - h \left(x^k \right)}{t} \\
&= \sup \left\{ \left\langle v, d^k \right\rangle, v \in \partial g \left(x^k \right) \right\} - \left\langle \nabla h \left(x^k \right), d^k \right\rangle \\
&= \sup \left\{ \left\langle v - \nabla h \left(x^k \right), d^k \right\rangle, v \in \partial g \left(x^k \right) \right\}.
\end{aligned} \tag{9}$$

根据定理 1 和式（8），

$$\left\langle \nabla h\left(x^k\right) - v, d^k \right\rangle \geq \rho \left\| d^k \right\|^2, \forall v \in \partial g\left(x^k\right) \tag{10}$$

将式（9）与式（10结合，即得第一个断言。

对于第二个断言，若 $d^k = 0$，则证明完成。否则，根据式（9）和第一个断言，可得：

$$\lim_{t \downarrow 0} \frac{\phi\left(x^k + t d^k\right) - \phi\left(x^k\right)}{t} \leq -\rho \left\| d^k \right\|^2 < -\alpha \left\| d^k \right\|^2.$$

因此，存在某个 $\sigma_k \geq 0$，使得：

$$\frac{\phi\left(x^k + \lambda d^k\right) - \phi\left(x^k\right)}{\lambda} \leq -\alpha \left\| d^k \right\|^2, \forall \lambda \in]0, \sigma_k].$$

□

如下结果表明，我们的 IBDCA 可视为经典 BDCA 的一种推广。

命题 6 *For any $k \in \mathcal{N}$, the following properties hold:*

1. Take $\alpha_k \leq \rho$, if $\phi'\ y^k\ d^k$ then there exists $\tau_k \geq 0$ such that for any $\lambda_k \in]1\ 1+\tau_k]$ eq. 6 and eq. 7 are satisfied.

(;) $\leq -\alpha_k \left\| d^k \right\|^2$, , , , () 0 *2. Take $\alpha_k \leq \rho$, if ϕ is smooth at point y^k, then there exists $\tau_k \geq 0$ such that for any , , , () 0 $\lambda_k \in]1\ 1+\tau_k]$ eq. 6 and eq. 7 are satisfied.*

Proof 根据方向导数的定义，我们有：

$$\lim_{t \downarrow 0} \frac{\phi\left(y^k + t d^k\right) - \phi\left(y^k\right)}{t} < -\alpha_k \left\| d^k \right\|^2,$$

因此，存在 $\tau_k \geq 0$，使得对任意 $\lambda \in]1, 1+\tau_k]$，我们有：

$$\phi\left(y^k + (\lambda - 1) d^k\right) - \phi\left(y^k\right) \leq -\alpha_k (\lambda - 1) \left\| d^k \right\|^2,$$

此外，我们有：

$$\phi\left(x^k + \lambda d^k\right) \leq \phi\left(y^k\right) - \alpha_k (\lambda - 1) \left\| d^k \right\|^2 \leq \phi\left(x^k\right) - \lambda \alpha_k \left\| d^k \right\|^2.$$

最后一个不等式由定理 4 和 $\alpha_k \leq \rho$.

根据定理 4 中等式 (8) 和 $\alpha_k < \rho$，我们有：

$$\phi'\left(y^k; d^k\right) = \left\langle \nabla \phi\left(y^k\right), d^k \right\rangle \leq -\rho \left\| d^k \right\|^2 < -\alpha_k \left\| d^k \right\|^2,$$

其余部分可用与第一个断言相同的方法证明。 □

接下来，我们给出 IBDCA 生成的序列的收敛结果。

定理 7 *For any initial point $x^0 \in \mathbb{R}^m$, IBDCA either returns a critical point or generates an infinite sequence that satisfies the following properties:*

1. *The sequence $\left\{\phi(x^k)\right\}$ is monotonically decreasing and converges to ϕ^*;*
2. *Any limit point of x^k is a critical point of the DC problem eq. (1). Furthermore, if ϕ is coercive, i.e., $\phi(x) \to +\infty$ as $\|x\| \to \infty$, then there exists a subsequence of x^k that converges to a critical point of the DC problem eq. (1);*
3. *$\sum_{k=0}^{\infty} \left\|d^k\right\|^2 < \infty$, and consequently, $\sum_{k=0}^{\infty} \left\|x^{k+1} - x^k\right\|^2 < \infty$.*

Proof 如果算法返回 x^k，根据定理 4， x^k 是一个临界点。否则，我们有：

$$\phi(x^{k+1}) \leq \phi(y^k) \leq \phi(x^k) - \rho\|d^k\|^2. \tag{11}$$

根据假设3，该序列单调递减且有下界，因此将收敛到 ϕ^*。

因此，我们有

$$\phi(x^k) - \phi(x^{k+1}) \to 0,$$

由此可得 $\|d^k\| \to 0$，根据式（11）。

令 x^* 为 x^k的任意一个临界点，并设 $\left\{x^{k_n}\right\}$ 是一个收敛到 x^*的子序列。由于 $\|y^{k_n} - x^{k_n}\|^2 = \|d^{k_n}\|^2 \to 0$，我们有 $y^{k_n} \to x^*$。根据连续性，可得：

$$\nabla h(y^{k_n}) \to \nabla h(x^*).$$

此外， $\nabla h(y^{k_n}) \in \partial g(x^{k_n})$，且 ∂g 的图像是闭的。因此，我们有

$$\nabla h(x^*) \in \partial g(x^*).$$

特别地，当 ϕ 是强制函数时，第一个断言意味着序列 $\{x^k\}$ 有界。因此，存在一个收敛到DC问题（1）的临界点的子序列。

对于最后一个断言，由于我们有：

$$\rho \left\|d^k\right\|^2 \leq \phi\left(x^k\right) - \phi\left(x^{k+1}\right).$$

然后，将这个不等式从 $k = 0$ 求和至 N，得到：

$$\sum_{k=0}^{N} \rho \left\|d^k\right\|^2 \leq \phi\left(x^0\right) - \phi\left(x^{N+1}\right) \leq \phi\left(x^0\right) - \inf_{x \in \mathbb{R}^m} \phi(x),$$

令 $N \to \infty$，即可完成证明。 □

最后，我们给出两个例子来展示 IBDCA 的优势。第一个例子见 [1,例 3.2], ，它表明，当 DC 分解中的函数 g 不可微时，即使 y^k 不是全局极小点， d^k 仍可能是 y^k 的上升方向。在这个例子中，BDCA 无法完成任务；nmBDCA 效率低下，因为它沿上升方向增加函数值；而改进的 BDCA 则能通过单调线搜索有效地解决该问题。

Example 3.1 我们考虑以下 DC 问题：

$$\min_{x=(u,v)\in\mathbb{R}^2} g(u,v) - h(u,v) := \phi(x),$$

其中

$$g(u,v) = -\frac{5}{2}u + u^2 + v^2 + |u| + |v|, \quad h(u,v) = \frac{1}{2}\left(u^2 + v^2\right),$$

设 $x^0 = (u^0, v^0) = (12, 1)$。那么，DCA生成的点为 $y^0 = (1, 0)$。如 [1, 示例 3.2], 所述，由于 $d^0 = y^0 - x^0 = (12, -1)$ 在点 y^0 处不是下降方向，因此无法执行 BDCA 的线搜索。

然而，我们改进的BDCA放弃了在 y^0 处的额外搜索。令 $x^1 = y^0$，从而通过DCA得到 $y^1 = (54, 0)$。随后，$x^2 = (\frac{3}{2}, 0)$ 是满足IBDCA搜索准则的点。事实上，取 $d^1 = y^1 - x^1 = (\frac{1}{4}, 0)$，经过简单计算可得：

$$\phi\left(x^2\right) = \phi\left(x^1 + 2d^1\right) < \phi\left(x^1\right) - 2\alpha\left\|d^1\right\|^2, \quad \forall \alpha \in]0, 2[.$$

此外，x^1 不是 ϕ的极小点，因此BDCA无法求解该模型。DCA生成的点收敛至全局最小值，而改进的BDCA的收敛速度显著加快，如图 1所示。

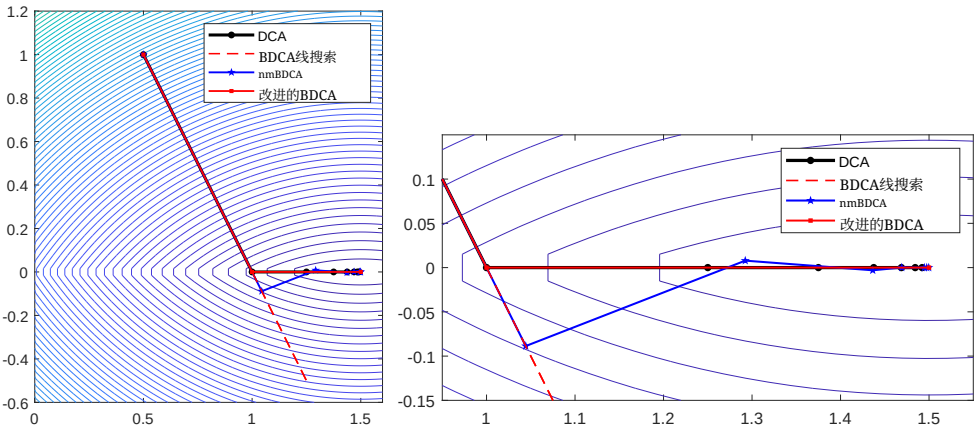


图 1 示例中 DCA、nmBDCA、BDCA 和 IBDCA 的迭代行为 3.1。

[1, 示例 3.1] 指出，BDCA 更有可能摆脱非全局解。下面这个改编自 SCAD（平滑裁剪绝对偏差，[40]）函数的示例表明，我们改进后的 BDCA 也具有这一优势。

Example 3.2 我们考虑以下 DC问题：

$$\min_{x=(u,v)\in\mathbb{R}^2} \phi(x) := \tilde{\phi}(u) + \tilde{\phi}(v),$$

其中

$$\tilde{\phi}(u) = \tilde{g}(u) - \tilde{h}(u),$$

$$\tilde{g}(u) = \begin{cases} |x| + \frac{x^2}{5} & , \text{ if } |x| < 2 \\ |x| + (|x| - 2)^2 + \frac{x^2}{5} & , \text{ if } |x| \geq 2 \end{cases},$$

$$\tilde{h}(u) = \begin{cases} \frac{x^2}{5} & , |x| \leq 1 \\ \frac{x^2 - 2|x| + 1}{2} + \frac{x^2}{5} & , 1 < |x| < 2. \\ \text{sign}(x) + \frac{x^2}{5} & , |x| > 2 \end{cases}$$

该DC问题满足?? 1–3。

图 2 展示了以初始点 $x^0 = (2.2, 0.4)$进行DCA、nmBDCA和改进的BDCA时的迭代行为。在改进的BDCA中，我们取 $\alpha_k = 0.2$, $\lambda_k = 3$, $\beta_k = 0.7$；对于nmBDCA，我们取 $\alpha_k = 0.2$, $\beta_k = 0.7$, $\lambda_k = 3 - 1$, $\nu_k = \left\|d^k\right\|^2 / (k + 1)$。我们的改进的BDCA不仅比DCA和nmBDCA具有更快的收敛速度，而且线搜索过程还避免了生成序列收敛到临界局部解(2，0)，因为该解甚至不是局部最小点。

我们在范围 $[0, 3] \times [0, 3]$ 内随机选取一百万个初始点。数值实验表明，DCA 生成的序列极有可能收敛到对应于非全局最小值的临界点，而 nmBDCA 仍有一部分点会收敛到对应于非全局最小值的临界点。相比之下，改进的BDCA 对所有选取的初始点均收敛到全局最小点，且收敛速度快于 nmBDCA。

	(2，2)	(0, 2)	(2, 0)	(0，0)	Time
DCA	110, 831 222 , 087 223 , 147 443 , 935				13.10 秒
nmBDCA	0	9645	9563	980792	276.7 秒
改进的BDCA	0	0	0	1，000，000	34。59 秒

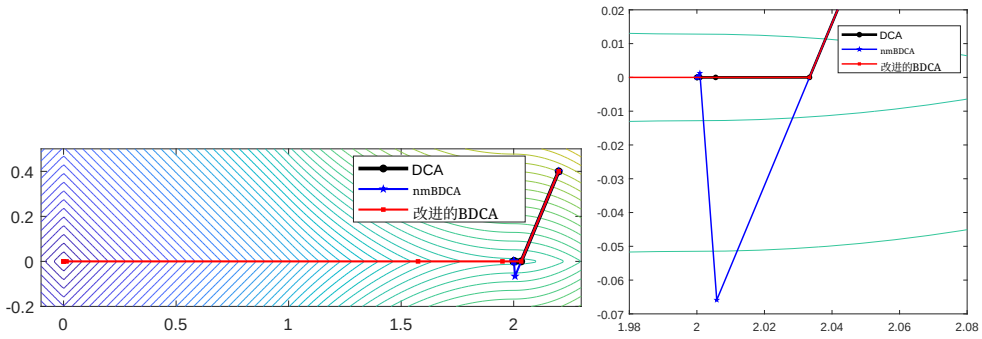


图2 DCA、nmBDCA和改进的BDCA以初始点 $x^0 = (2.2, 0.4)$在示例3.2中的迭代行为。

4 基于 Kurdyka–-Lojasiewicz 不等式的全局收敛性质

本节基于 Kurdyka– -Lojasiewicz (K-L) 不等式，建立改进的BDCA 所生成序列的全局收敛性和收敛速率。首先，我们回顾 K-L 不等式的相关概念，该不等式已广泛应用于差凸（DC）问题和非光滑优化[1, 18, 41–43]。

定义 6 设 $f:\mathbb{R}^m \to \mathbb{R}$ 是一个局部利普希茨连续函数。若存在 $\eta \in]0, +\infty[$，一个 x^* 的邻域 U，以及一个凹函数 $\psi:[0,\eta] \to [0,+\infty[$，使得 f 在 $x^* \in \mathbb{R}^m$ 处满足 K-L 不等式，则称 f 在 $x^* \in \mathbb{R}^m$ 处满足 K-L 不等式：

1. $\psi(0) = 0$；2. ψ 在区间 $[0,\eta]$ 上属于 C^1 类；3. $\psi'[0,\eta] > 0$；
4. 对任意满足 $f(x^*) \le f(x) \le f(x^*) + \eta$ 的 $x \in U$，存在：

$$\psi'\left(f(x) - f\left(x^*\right)\right) \operatorname{dist}\left(0, \partial_C f(x)\right) \ge 1 \tag{12}$$

接着，我们给出次解析函数和半代数函数的定义（参见 [44–46]）。这类函数涵盖了优化问题中常见的广泛函数类别。在一定条件下，这些函数满足 K-L 不等式。

定义 7 我们称 $A \subset \mathbb{R}^m$ 为次解析的，如果对于任意点 $a \in A$，存在 a 的一个邻域 V，以及某个 $d \ge 1$，使得 $A \cap V = \{x \in \mathbb{R}^m : \exists y \in \mathbb{R}^d, (x, y) \in B\}$，其中 B 有界且满足：

$$B = \bigcup_{i=1}^{p} \bigcap_{j=1}^{q} \left\{(x,y) \in \mathbb{R}^m \times \mathbb{R}^d : f_{ij}(x,y) = 0, g_{ij}(x,y) < 0\right\},$$

对于某些实解析函数 $f_{ij}, g_{ij}: \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R} (1 \le i \le p, 1 \le j \le q)$。

函数 $f: \mathbb{R}^m \to \mathbb{R} \cup +\infty$ $f = \{x\ t \quad x \in :$ 如果其图 gph (,): , $f() = t\}$ $\mathbb{R}^m$ 是 的 x 一个次解析集，则称其为次解析函数。$\mathbb{R}^m \times \mathbb{R}$

定义 8 我们称集合 $A \subset \mathbb{R}^m$ 是半代数的，如果 A 可以表示为

$$A = \bigcup_{i=1}^{p} \bigcap_{j=1}^{q} \{x \in V : f_{ij}(x) = 0, g_{ij}(x) < 0\},$$

其中 $f_{ij}, g_{ij}: \mathbb{R}^m \to \mathbb{R} (1 \le i \le p, 1 \le j \le q)$ 是实多项式函数。

以下是优化问题中常见次解析函数的一些示例（参见 [44–46]）。

命题 8 *The following conclusions hold:*

1. The following functions are subanalytic:
(a) Real-analytic function; (b) Semialgebraic function; (c) The sum of a real-analytic function and a subanalytic function.

2. The following functions are semialgebraic:
(a) The finite sum and product of semialgebraic functions; (b) The composition of semialgebraic functions; (c) l_p norm, i.e., $\|x\|_p := (\sum_i |x^i|^p)^{1/p}$, $p \in \mathbb{N}^+$.

下面的定理在全局收敛分析中起着关键作用。

定理 9 *([47, Theorem 3.1]): Let $f \ \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ be a subanalytic function $domf:$,, be closed and $f|_{Domf}$ be continuous. Let x^* be a critical point of f. Then f satisfies the () , Kurdyka-Łojasiewicz inequality eq. 12 at x^*. Besides the concave function ϕ as in the Kurdyka-Łojasiewicz inequality can be chosen as ϕ t ()*

$= Mt^{1-\theta}$ for some $M > 0$ and $\theta \in [0, 1[$, where we use the convention $0^0 = 1$. θ is called the Łojasiewicz exponent.

接下来，我们应用上述结果，建立 IBDCA 所生成序列的全局收敛定理。

定理 10 *Let the DC function h be subanalytic and $\phi|_{Domf}$ be continuous. Suppose that h is coercive and that the sequence $\{x^k\}$ generated by the IBDCA has a subsequence converging to a critical point x^* (by theorem 7).*

Assume for any $k \in \mathbb{N}^+$ we have $\lambda_k \leq \bar{\lambda}$ for some $\bar{\lambda} > 1$ then the sequence $\left\{x^k\right\}$ converges ,,, to x^*.

Proof 令 X 为 $\left\{x^k\right\}$的聚点集合。根据定理7，有 $\|d^k\| \to 0$,

$\|x^{k+1} - x^k\| \to 0 \quad k \to \infty \qquad x \in X \ \ \phi\ x \ \equiv \phi^* = \phi\ x^*\ . \qquad \|d^k\| =$ 当时
$\|y^k - x^k\| \to 0 \qquad X \qquad\qquad\qquad\

取 $\theta = \max_i \{\theta_{z^i}\}, K = \min_i \{(1-\theta_{z^i})M_{z^i}\}$ 以及 $L = \min_i \left\{L_{z^i}\right\}$，并利用对任意 z^i均有 $\phi(z^i) = \phi(x^*)$这一事实，则我们有：

$$\left|\phi\left(y^k\right) - \phi\left(x^*\right)\right|^\theta \leq K \operatorname{dist}\left(0, \partial_C\left(y^k\right)\right), \tag{13}$$

and

$$\left\|\nabla h\left(y^k\right) - \nabla h\left(x^k\right)\right\| \leq L\left\|y^k - x^k\right\| = L\left\|d^k\right\|. \tag{14}$$

利用式 (8)，可得

$$\nabla h\left(x^k\right) - \nabla h\left(y^k\right) \in \partial g\left(y^k\right) - \nabla h\left(y^k\right) = \partial_C\left(\phi\left(y^k\right)\right). \tag{15}$$

结合式 (13) 至 (15)，可得：

$$\left|\phi\left(y^k\right) - \phi\left(x^*\right)\right|^\theta \leq KL\left\|d^k\right\|. \tag{16}$$

□

由于 $\phi(t) = t^{1-\theta}$ 在 $[0, +\infty[$, 处为凹函数，根据定理 4，可得：

$$\begin{aligned}
&\left(\phi\left(y^k\right) - \phi\left(x^*\right)\right)^{1-\theta} - \left(\phi\left(y^{k+1}\right) - \phi\left(x^*\right)\right)^{1-\theta} \\
\geq&(1-\theta)\left(\phi\left(y^k\right) - \phi\left(x^*\right)\right)^{-\theta}\left(\phi\left(y^k\right) - \phi\left(y^{k+1}\right)\right) \\
\geq&(1-\theta)\left(\phi\left(y^k\right) - \phi\left(x^*\right)\right)^{-\theta}\left(\phi\left(x^{k+1}\right) - \phi\left(y^{k+1}\right)\right) \\
\geq&(1-\theta)\left(\phi\left(y^k\right) - \phi\left(x^*\right)\right)^{-\theta}\rho\left\|d^{k+1}\right\|^2.
\end{aligned} \tag{17}$$

结合式 (16) 和 (17)，并应用不等式 $\frac{a^2}{b} \leq 2a - b, \ \forall a, \ b \succ_- 0$，可得

$$\begin{aligned}
2\left\|d^{k+1}\right\| - \left\|d^k\right\| &\leq \frac{\left\|d^{k+1}\right\|^2}{\left\|d^k\right\|} + \left\|d^k\right\| \\
&\leq \frac{KL}{(1-\theta)\rho}\left(\left(\phi\left(y^k\right) - \phi\left(x^*\right)\right)^{1-\theta} - \left(\phi\left(y^{k+1}\right) - \phi\left(x^*\right)\right)^{1-\theta}\right),
\end{aligned} \tag{18}$$

将上述不等式从 0 到 N 求和，并令 $N \to \infty$，可得：

$$\sum_{k=1}^{\infty}\left\|d^k\right\| \leq \left\|d^0\right\| + \frac{KL}{(1-\theta)\rho}\left(\phi\left(y^k\right) - \phi\left(x^*\right)\right)^{1-\theta}.$$

由于 $\lambda_k \leq \lambda$，因此：

$$\sum_{k=1}^{\infty}\left\|x^{k+1} - y^k\right\| < (\bar{\lambda} - 1)\sum_{k=1}^{\infty}\left\|d^k\right\| < \infty.$$

因此，$\left\{\sum_{k=1}^{N}\left\|d^k\right\|\right\}$ 和 $\left\{\sum_{k=1}^{N}\left\|x^{k+1}-y^k\right\|\right\}$ 都是柯西序列。对于任意整数 N 和 $p\in\mathbb{N}^+$,

$$\left\|x^N-x^{N+p}\right\|\leq\sum_{k=N}^{N+p-1}\left\|x^{k+1}-x^k\right\|\leq\sum_{k=N}^{N+p-1}\left\|x^{k+1}-y^k\right\|+\sum_{k=N}^{N+p-1}\left\|d^k\right\|.$$

因此，$\left\{x^k\right\}$ 也是柯西序列，$x^k\to x^*$，因为 $k\to\infty$。

Remark 2 在实践中，我们通常令 $\lambda_k=\lambda>1,\ \forall k$，这种方法在以往研究中被广泛应用 [18, 29]。

定理

11 *Assume the assumption in theorem 10 holds. Let x^* be the limit point of $\{x^k\}$ and denote the Łojasiewicz exponent in theorem 10 as $\theta\in[0,1[$.*

1. *If $\theta=0$, then sequence $\{x^k\}$ converges to x^* after a finite number of iterations.*
2. *If $\theta\in]0,\frac{1}{2}]$, then*

$$\left\|x^k-x^*\right\|\leq c_1\left(\frac{\gamma}{1+\gamma}\right)^k,$$

where $\gamma=1+\frac{K^{\frac{1}{\theta}}L^{\frac{1}{\theta}}}{(1-\theta)\rho}, c_1=\bar{\lambda}\sum_{k=0}^{\infty}\left\|d^k\right\|$;

3. *If $\theta\in]\frac{1}{2},1[$, then:*

$$\left\|x^k-x^*\right\|\leq c_2k^{\frac{1-\theta}{1-2\theta}},$$

where $c_2=\bar{\lambda}\left(\frac{2\theta-1}{\gamma(1-\theta)}\right)^{\frac{1-\theta}{1-2\theta}}$.

Proof 如果 $\theta=0$，根据式（16），我们有:

$$\left\|d^k\right\|\geq\frac{1}{KL},$$

这给出

$$f\left(x^k\right)-f\left(x^{k+1}\right)\geq\rho\left\|d^k\right\|^2\geq\frac{\rho}{K^2L^2}.$$

因此，该序列 $\{x^k\}$ 在有限次迭代后收敛于 x^*。

如果 $0<\theta<1$，则设 $D_k=\sum_{j=k}^{\infty}\|d^k\|$。对于 $k>0,\ p>0$，将式（18）从 $k-1$ 到$k+p$ 求和，并令 $p\to\infty$，可得

$$D_k\leq\left\|d^{k-1}\right\|+\frac{KL}{(1-\theta)\rho}\left(\phi\left(y^{k-1}\right)-\phi\left(x^*\right)\right)^{1-\theta}.\tag{19}$$

□

结合式（16）与式（19）可得

$$D_k\leq\left\|d^{k-1}\right\|+\frac{K^{\frac{1}{\theta}}L^{\frac{1}{\theta}}}{(1-\theta)\rho}\left\|d^{k-1}\right\|^{\frac{1-\theta}{\theta}},\tag{20}$$

其中，不失一般性，我们假设对任意 k, $\|d^k\| < 1$。

如果 $\theta \in]0, \frac{1}{2}]$，那么 $\|d^k\|^{1-\theta} \leq \|d^k\|$。因此，我们可以由式（20）得到：

$$D_k \leq \gamma \left(D_{k-1} - D_k\right)^{\frac{1-\theta}{\theta}}.$$

令 $\gamma = 1 + \frac{k\theta}{L\rho(1-\theta)\rho}$，我们得到

$$D_k^{\frac{1-2\theta}{1-\theta}} - D_{k-1}^{\frac{1-2\theta}{1-\theta}} \geq \frac{1-2\theta}{1-\theta} D_k^{\frac{-\theta}{1-\theta}} \left(D_k - D_{k-1}\right) \geq \frac{2\theta - 1}{\gamma(1-\theta)}.$$

将这些不等式累加至 ∞ 并利用 $\lambda_j \leq \lambda$，于是得到：

$$\left\|x^k - x^*\right\| \leq \bar{\lambda} D_k \leq \bar{\lambda} \left(D_0^{\frac{1-2\theta}{1-\theta}} + \frac{2\theta - 1}{\gamma(1-\theta)} k\right)^{\frac{1-\theta}{1-2\theta}} \leq \bar{\lambda} \left(\frac{2\theta - 1}{\gamma(1-\theta)} k\right)^{\frac{1-\theta}{1-2\theta}}.$$

5 图像复原中的应用

图像去噪是低级图像处理的重要组成部分 [48, 49]。尽管已有许多数学变分方法专门用于处理加性高斯噪声，但由于非高斯噪声在各种科学和工程应用中普遍存在，对其的研究仍有很大的发展潜力。本节介绍如何应用所提出的 IBDCA，高效求解图像复原中的一个重要 DC问题。

5.1 复原受柯西噪声污染的退化图像

柯西噪声是一种脉冲噪声。数学上，含有柯西噪声的退化图像可表示为 $f = u + v$，其中 $u \in \mathbb{R}^m = \mathbb{R}^{m_1 \times m_2}$ 是原始图像，$v \in \mathbb{R}^m$ 表示某种柯西噪声。若随机变量 V 的概率密度函数如下，则称其服从（中心）柯西分布：

$$g(v) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + v^2}.$$

基于最大后验估计（MAP）的柯西噪声去除模型[20] 可表述如下：

$$\min_{u \in R^m} E(u) := TV(u) + \frac{\mu}{2} \left\langle \log\left(\gamma^2 + (u - f)^2\right), \mathbf{1} \right\rangle, \tag{21}$$

其中 $TV(u) = \|\nabla u\|_{2,1} = \left\| \sqrt{|\nabla_x u|^2 + |\nabla_y u|^2} \right\|_1$ 表示离散全变差[50]。表示水平方向和垂直方向的梯度算子分别记为 ∇_x 和 ∇_y，而 $\mu > 0$ 是全变差正则化项与保真项之间的权衡参数。

在 [51], 采用非凸交替方向乘子法（ADMM）求解式（21），而这里我们使用DC方法。式（21）可表示为一个 DC问题：

$$\min_{u\in\mathbb{R}^m} G(u) - H(u),$$

其中：

$$\begin{aligned} G(u) &= TV(u) + \frac{c}{2}\|u\|^2, \\ H(u) &= -\frac{\mu}{2}\left\langle \log\left(\gamma^2 + (u-f)^2\right), \mathbf{1}\right\rangle + \frac{c}{2}\|u\|^2, \end{aligned} \tag{22}$$

这里， $c>0$ 表示确保 $G(u)$和 $H(u)$具有强凸性的参数。首先，我们给出以下引理。

命题 12 *When $c \geq \mu/\gamma^2$, the function H defined in eq. (22 when) is convex. Thus,,() $c > \mu/\gamma^2$ H is $c-\mu/\gamma^2$ -strongly convex.*

Proof 只需考虑单变量情形。令

$$h(t) = -\frac{\mu}{2}\log\left(\gamma^2 + t^2\right) + \frac{c}{2}t^2.$$

然后，通过直接计算可得：

$$\begin{aligned} h'(t) &= -\frac{\mu t}{\gamma^2 + t^2} + ct, \\ h''(t) &= \frac{\mu(t^2-\gamma^2)}{(\gamma^2+t^2)^2} + c. \end{aligned}$$

我们只需确保 $h''(t) \leq 0$ ，使得 $h(t)$为凹函数，即：

$$c \geq \frac{\mu(\gamma^2 - t^2)}{(\gamma^2+t^2)^2}.$$

请注意，上述不等式右侧在 $t=0$ 处达到最大值。 □

求解方程（21）的 DC 过程如下：

$$\begin{aligned} v^k &= \nabla h\left(u^k\right) = \frac{\mu\left(u^k - f\right)}{\gamma^2 + \left(u^k - f\right)^2} + cu^k. \\ u^k &= \arg\min_u TV(u) + \frac{c}{2}\|u\|^2 - \left\langle v^k, u\right\rangle. \end{aligned}$$

[52] 中的原始-对偶算法可用于高效求解 u-子问题。由于 u-子问题中的目标函数是强凸的，原始-对偶算法的收敛速率可加快至 $O(1/N^2)$ [52]。

根据定理8，DC问题（21）中的目标函数是次解析的。此外， H 是局部利普希茨连续的，且 $E = G - H$ 是强制的。因此，定理10和11保证了所提算法的全局收敛。



(b) 房屋



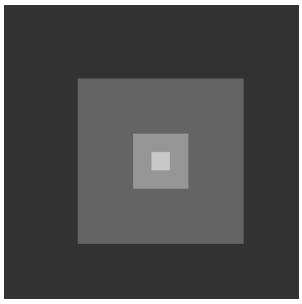
(c) 蝴蝶



(d) 男人



(e) Bird



(f) 平方

图 3 测试图像：摄影师、房屋、蝴蝶、男人、鸟和平方。在这些图像中，摄影师、房屋、蝴蝶、鸟和平方的尺寸为 256×256。男人的尺寸为 512×512。

5.2 数值实验

本节将我们的框架应用于柯西噪声下的图像去噪任务，以检验所提算法的有效性。测试中，我们采用3所示的五幅不同尺寸的测试图像作为测试集。按照如下方式，根据柯西分布添加柯西噪声以生成噪声图像：

$$f = u + \gamma \frac{v_1}{v_2},$$

其中，随机变量 v_1、 v_2 服从正态分布 $\mathcal{N}(0,\ 1)$。为改变柯西噪声水平，我们分别令 γ 为 3 和 5。 γ 越大，图像看起来越嘈杂，图像复原也就越困难。

我们将提出的 IBDCA 与三种具有竞争力的方法进行比较，包括 ADMM [51], DCA [53], 和 nmBDCA [39]。由于文献 [39], 中的数值实验表明，与某些 PPMDC（用于差分凸函数的邻近点方法，[15, 54]）相比，nmBDCA 的收敛速度更快、精度更高，因此本文不与这些方法进行比较。

为确保比较公平，我们对各种竞争方法进行了精确调整，并采用其最优参数配置进行设置，从而获得最令人满意的成像结果。这使我们能够全面且客观地评估不同技术的相对性能。

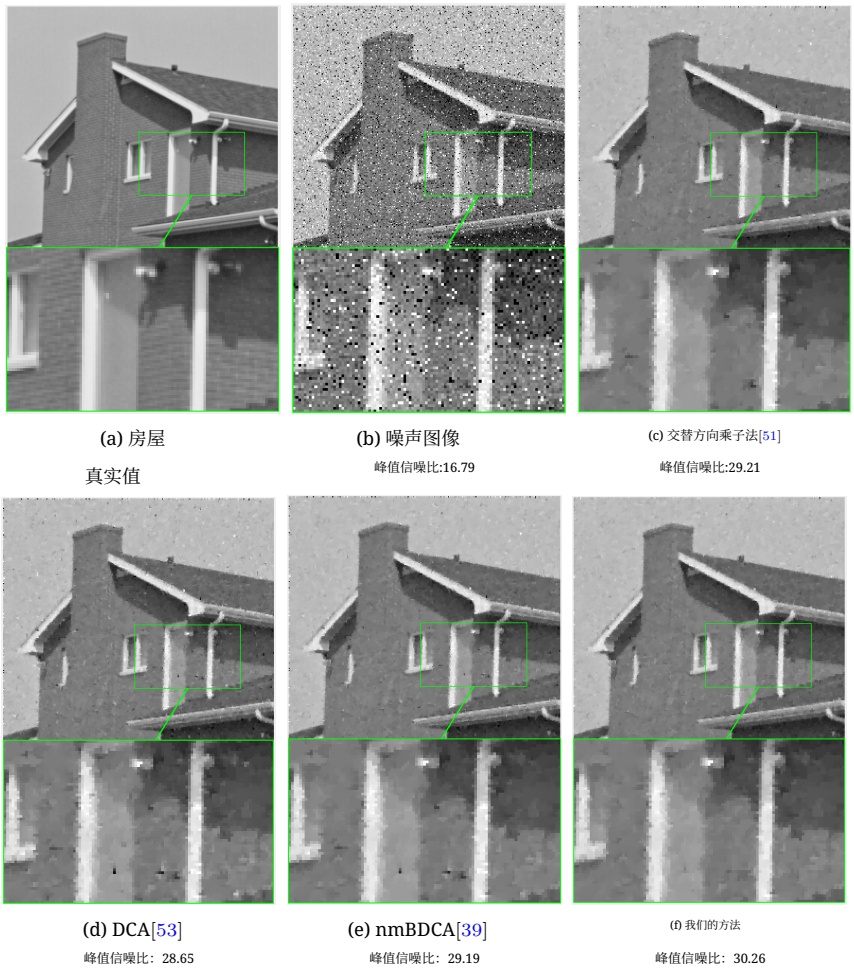


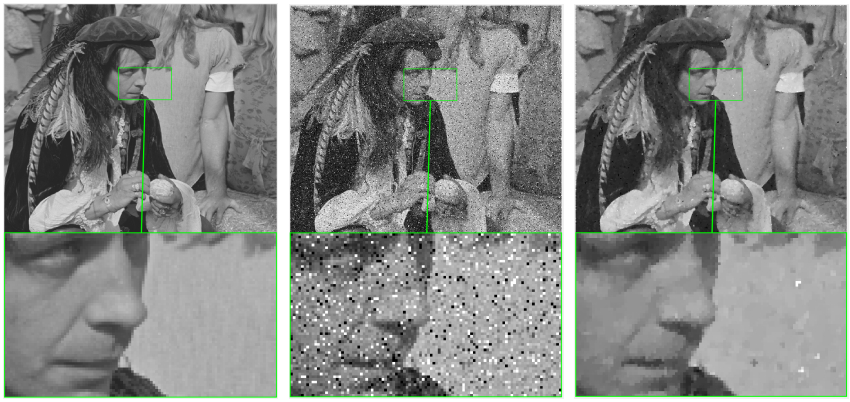
图 4 对所选复原图像的视觉比较，该图像含有 $\gamma = 3$的柯西噪声。放大以获得更好的可视化效果。

为评估数值指标，我们采用峰值信噪比（PSNR）和相对误差（ReErr）来衡量去噪结果的质量。峰值信噪比定义如下：

$$\mathrm{PSNR} = 20\log_{10}\frac{255\sqrt{m_1 m_2}}{\|u^\star - u\|_F}, \tag{23}$$

其中，$\|\cdot\|_F$ 表示弗罗贝尼乌斯范数，u 表示原始图像，$u^\star$ 表示复原图像。m_1 和 m_2 表示输入图像的尺寸。相对误差定义如下：

$$\mathrm{ReErr} = \frac{\|u^\star - u\|_2^2}{\|u\|_2^2}. \tag{24}$$



(a) 男人

真实值

(b) 噪声图像

峰值信噪比：16.74

(c) 交替方向乘子法[51]

峰值信噪比：28.09



(d) DCA[53]

峰值信噪比:27.79

(e) nmBDCA[39]

峰值信噪比:28.14

(f) 我们的方法

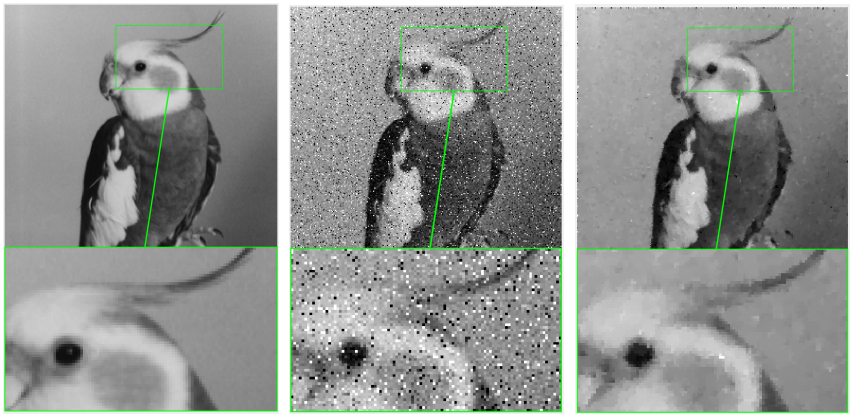
峰值信噪比:28.71

图 5 所选复原图像的视觉效果对比，该图像受到含有 $\gamma = 3$ 的柯西噪声污染。放大以获得更好的视觉效果。

所提算法的相关参数会影响不同柯西噪声强度下的实验结果。当 $\gamma = 3$时，将用于衡量保真项与正则化项之间权衡的正则化参数调整为 $\mu = 15$，并将所有方法的初始点设为 $u_0 = f$。最大迭代次数设为 200，为了获得更好的去噪结果，停止准则中的 tol 设为 5×10^{-4}：

$$\frac{\left|E\left(u^k\right) - E\left(u^{k+1}\right)\right|}{E\left(u^k\right)} \leq tol.$$

对于 $\gamma = 5$，我们调整 $\mu = 20$，其他参数和容差保持不变。值得注意的是，我们需要在任何情况下保证 $c \geq \mu/\gamma^2$。四种方法的具体参数如下：



(a) Bird

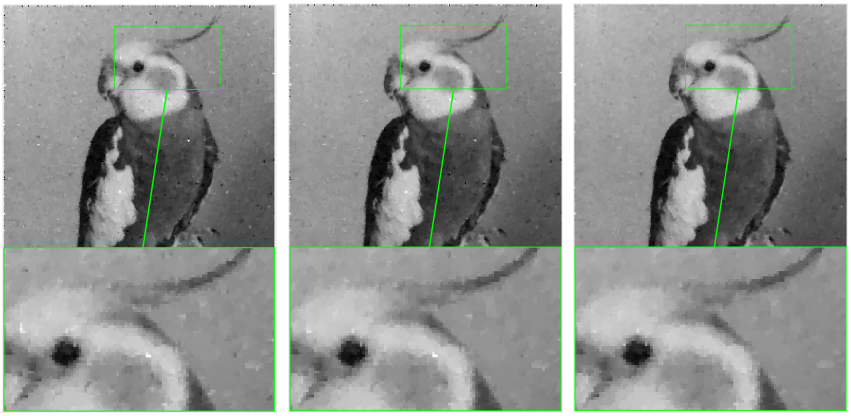
真实值

(b) 噪声图像

峰值信噪比：16.77

(c) 交替方向乘子法[51]

峰值信噪比：30.53



(d) DCA[53]

峰值信噪比：29.55

(e) nmBDCA[39]

峰值信噪比：30.15

(f) 我们的方法

峰值信噪比：31.75

图 6 所选复原图像的视觉对比，该图像受到柯西噪声影响，参数为 $\gamma = 3$。放大查看效果更佳。

1. 对于交替方向乘子法，惩罚项的系数应大于 μ/γ^2。因此，在数值实验中，我们将该参数设为 1。交替方向乘子法中的两个子问题分别通过原始-对偶算法和牛顿法 [51] 求解。2. 对于 DCA 、 nmBDCA 和我们提出的 IBDCA，在噪声水平为 $\gamma = 5$ 时，我们将参数 $c = 1.10$ 设为相应值。对于柯西噪声水平为 $\gamma = 3$ 的情况，为了获得更好的结果，我们将该值调整为 $c = 1.83$ 。3. 在 IBDCA 中，对于任意 k，线搜索的步长初始化为 $\lambda_k = 10$，超线搜索参数为 $\beta = 0.5$，并且 $\alpha_k = 0.9(c - \frac{\mu}{\gamma^2})$ 。4. 在 nmBDCA 中，对于任意 k，线搜索的步长初始化为 $\lambda_k = 9$（因为线搜索从 $u^k + d^k$ 开始），线搜索参数为 $\beta = 0.5$，$\alpha_k = 0.9(c - \frac{\mu}{\gamma^2})$，以及 $v_k = \frac{\|d^k\|^2}{k+1}$（[39], , 例 4.6）。

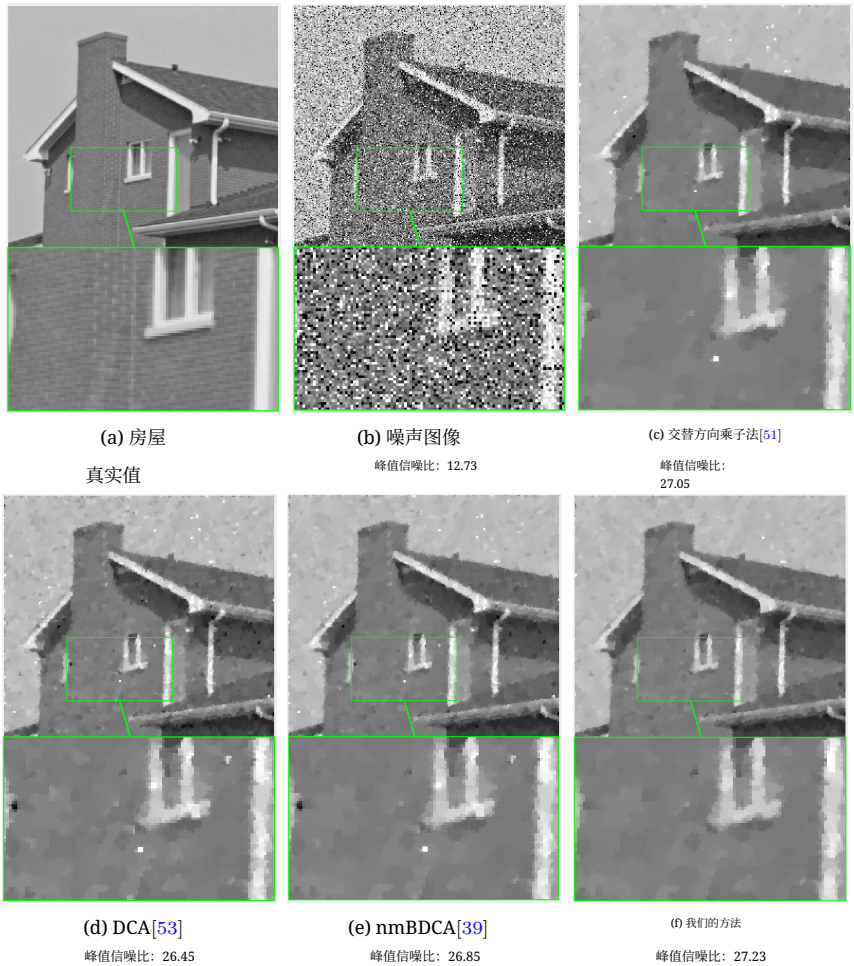


图 7 所选复原图像的视觉效果对比，该图像受到柯西噪声影响，噪声参数为 $\gamma = 5$。放大以获得更好的视觉效果。

表1 展示了在强度为 $\gamma = 3$,5 的柯西噪声下，我们对五幅测试图像进行去噪任务的详细结果。我们列出了这些结果对应的峰值信噪比（dB）、相对误差、CPU时间和迭代次数。从表中可以看出，与其他方法相比，我们提出的方法在峰值信噪比和相对误差方面表现更优，同时速度更快、收敛所需的迭代次数更少。我们还将三幅选定图像与其他方法进行了视觉对比，其中图4、图5和图6 展示了在强度为 $\gamma = 3$的柯西噪声下的复原图像；图7、图8 和图9展示了在强度为 $\gamma = 5$的柯西噪声下的去噪结果。视觉对比表明，我们的 IBDCA 方法始终能取得最佳视觉效果，生成更清晰的图像，并增强结构细节、减少噪声，这一点在绿色放大区域中尤为明显。

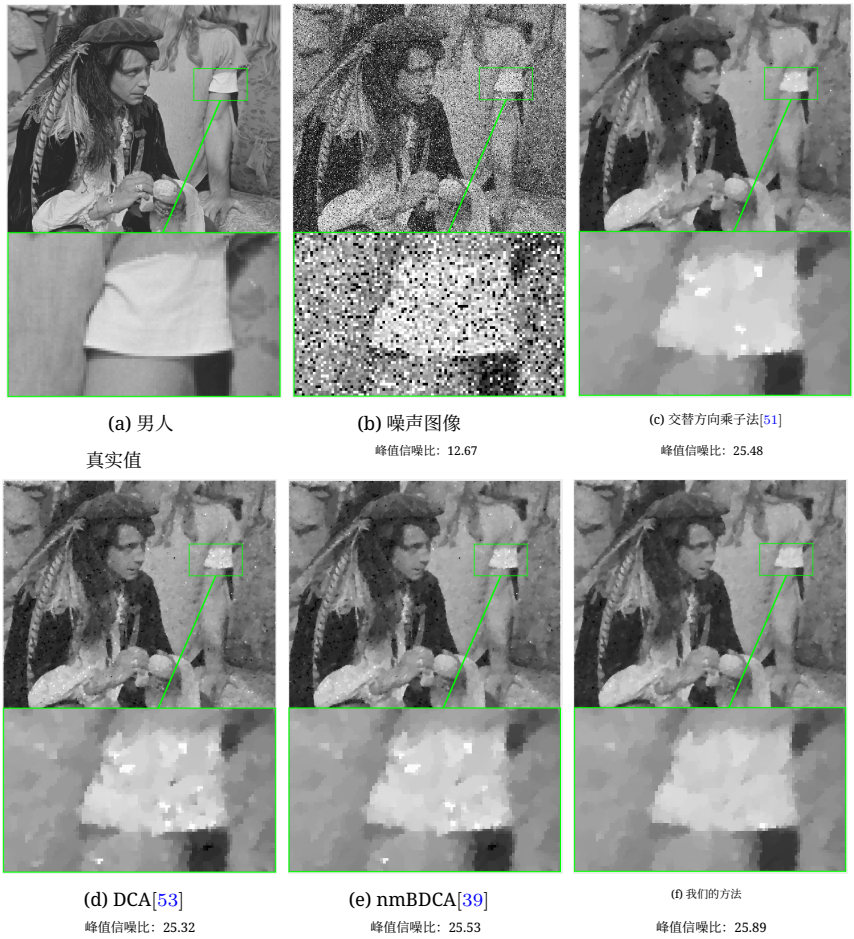


图8 对所选复原图像的视觉比较，该图像含有 $\gamma = 5$的柯西噪声。放大以获得更好的可视化效果。

图10 和11 展示了我们提出的方法及三种对比方法的能量对数和峰值信噪比值随CPU时间变化的关系。这些图示突显了我们提出的 IBDCA 方法在不同柯西噪声强度下求解 DC问题的优势。图中显示，我们提出的 IBDCA 方法能以最短的 CPU时间达到最低能量和最高峰值信噪比。

6 结论

我们针对一类常见的非光滑 DC 问题，创新性地提出了一种改进的增强型 DC 算法（IBDCA）。尽管对于这些问题，BDCA 在 [18] 和 [1] 中的搜索方向在当前迭代处不一定是下降方向，但在前一次迭代处是下降方向。基于这一观察，我们提出了一种线搜索策略，并证明在许多情况下，我们的

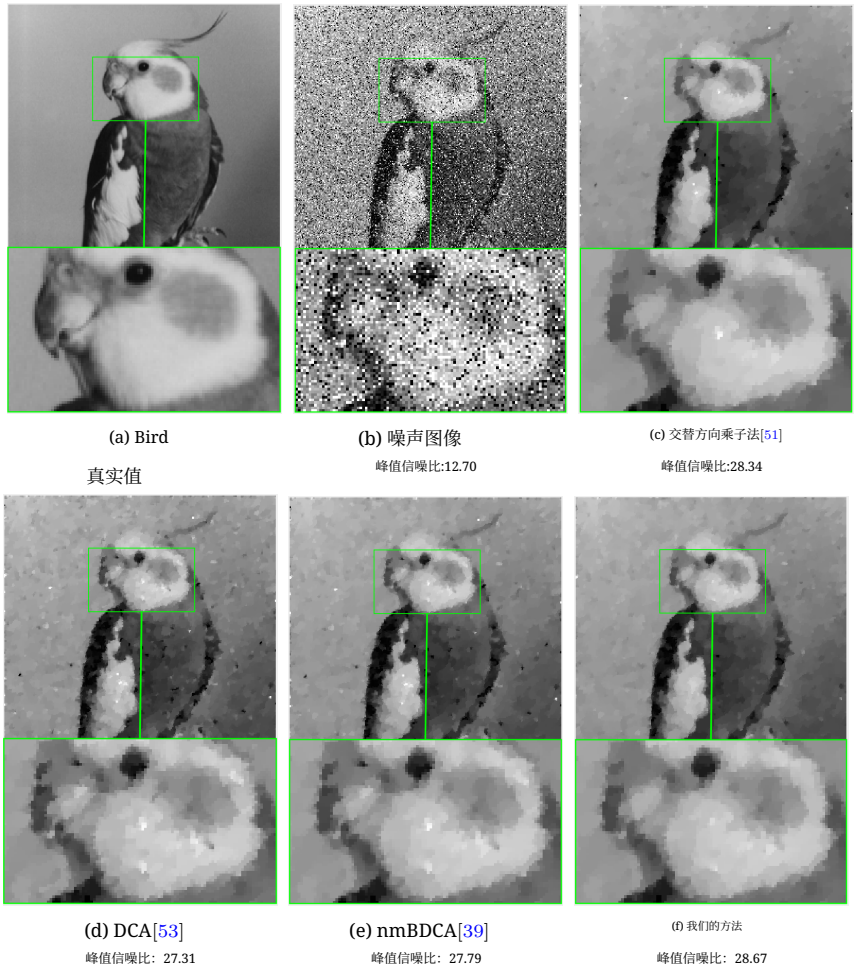


图 9 所选复原图像的视觉对比，该图像含有带 $\gamma = 5$ 的柯西噪声。放大以获得更好的视觉效果。

所提出的 IBDCA 只需少量额外计算，就能比 DCA 获得更大的下降幅度。具体而言，如果目标函数在当前迭代处光滑，IBDCA 就能实现与 BDCA 在 [18], 中类似的加速效果。因此，IBDCA 可视为 BDCA 的一种推广。在某些标准条件下，我们证明了 IBDCA 生成的每个聚点都是目标函数的临界点，并建立了新的单调下降性质。利用次解析函数的 Kurdyka-Łojasiewicz 性质，我们证明了全局收敛性和收敛速率。未来的研究可关注满足 Kurdyka-Łojasiewicz 不等式的一般函数是否也具有类似的收敛性质。数值实验和理论分析表明，IBDCA 的收敛速率更快，并能有效避免算法陷入局部次优解。此外，我们的线

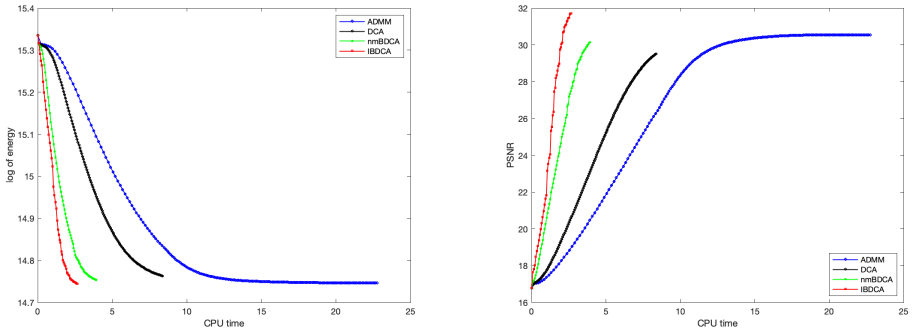


图10 所选图像“bird”在柯西噪声水平为 $\gamma = 3$ 时，我们提出的方法与对比方法的能量和峰值信噪比曲线。左：能量随时间变化图；右：峰值信噪比随时间变化图。可以看出，我们提出的方法显著优于其他三种对比方法。

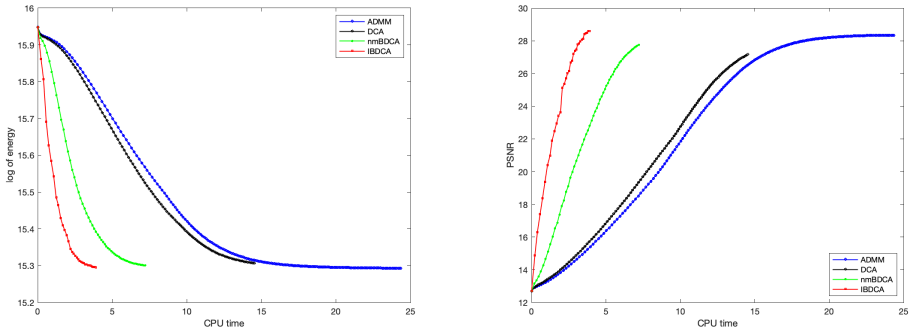


图11 所选图像“bird”在柯西噪声水平为 $\gamma = 5$ 时，我们提出的方法与对比方法的能量和峰值信噪比曲线。左：能量随时间变化图；右：峰值信噪比随时间变化图。可以看出，我们提出的方法显著优于其他三种对比方法。

搜索策略可以防止目标函数在某些潜在的非下降方向上增大，因此比 nmBDCA 更高效。探索将 IBDCA 应用于其他领域以解决非光滑或非凸问题，是未来研究的一个有前景的方向。

致谢

本研究部分经费来自国家重点研发计划（项目编号：2021YFE0203700）、ITF 项目（编号：MHP/038/20）、CRF 项目（编号：8730063）、RGC 项目（编号：14300219、14302920、14301121）以及香港中文大学研究直接拨款。

数据可用性

有关数据可用性的咨询，请联系作者。

利益冲突

作者声明，他们不存在利益冲突。

表 1 五幅测试图像去噪的峰值信噪比（dB）/相对误差值。柯西噪声水平分别为 $\gamma = 3,5$。文中还给出了重建图像所需的 CPU 时间（s）和迭代次数，并提供了各项指标的平均值。总体而言，在这两个不同的柯西噪声水平下，所提算法在这些对比方法中表现最佳。

图像	柯西噪声水平	方法	PSNR	相对误差	时间（s）	迭代次数
房屋	$\gamma = 3$	ADMM	29.21	0.0607	26.35	200
房屋	$\gamma = 3$	DCA	28.65	0.0648	9.59	200
房屋	$\gamma = 3$	nmBDCA	29.19	0.0609	4.64	71
房屋	$\gamma = 3$	Ours	**30.26**	**0.0538**	**2.93**	**45**
房屋	$\gamma = 5$	ADMM	27.05	0.0779	23.97	200
房屋	$\gamma = 5$	DCA	26.45	0.0834	15.15	122
房屋	$\gamma = 5$	nmBDCA	26.85	0.0797	6.99	61
房屋	$\gamma = 5$	Ours	**27.23**	**0.0763**	**3.63**	**29**
蝴蝶	$\gamma = 3$	ADMM	26.91	0.0923	33.87	200
蝴蝶	$\gamma = 3$	DCA	26.61	0.0955	10.34	146
蝴蝶	$\gamma = 3$	nmBDCA	26.93	0.0920	4.39	68
蝴蝶	$\gamma = 3$	Ours	**27.10**	**0.0913**	**3.37**	**51**
蝴蝶	$\gamma = 5$	ADMM	23.56	0.1356	28.50	200
蝴蝶	$\gamma = 5$	DCA	23.62	0.1347	17.71	121
蝴蝶	$\gamma = 5$	nmBDCA	23.79	0.1323	7.70	63
蝴蝶	$\gamma = 5$	Ours	**23.83**	**0.1319**	**3.68**	**30**
Man	$\gamma = 3$	ADMM	28.09	0.0827	117.43	200
Man	$\gamma = 3$	DCA	27.79	0.0857	37.40	154
Man	$\gamma = 3$	nmBDCA	28.14	0.0823	17.59	71
Man	$\gamma = 3$	Ours	**28.71**	**0.0771**	**11.72**	**46**
Man	$\gamma = 5$	ADMM	25.48	0.1118	111.38	200
Man	$\gamma = 5$	DCA	25.32	0.1138	66.55	123
Man	$\gamma = 5$	nmBDCA	25.53	0.1111	32.32	63
Man	$\gamma = 5$	Ours	**25.89**	**0.1066**	**16.41**	**30**
Bird	$\gamma = 3$	ADMM	30.53	0.0568	22.07	200
Bird	$\gamma = 3$	DCA	29.55	0.0635	8.21	150
Bird	$\gamma = 3$	nmBDCA	30.15	0.0593	3.88	69
Bird	$\gamma = 3$	Ours	**31.75**	**0.0493**	**2.64**	**46**
Bird	$\gamma = 5$	ADMM	28.34	0.0731	24.52	200
Bird	$\gamma = 5$	DCA	27.21	0.0833	14.47	119
Bird	$\gamma = 5$	nmBDCA	27.79	0.0778	7.03	60
Bird	$\gamma = 5$	Ours	**28.67**	**0.0703**	**3.94**	**32**
平方	$\gamma = 3$	ADMM	33.04	0.0774	30.43	200
平方	$\gamma = 3$	DCA	31.10	0.0968	30.43	157
平方	$\gamma = 3$	nmBDCA	**32.22**	**0.0852**	5.02	73
平方	$\gamma = 3$	Ours	32.18	0.0856	**4.47**	53
平方	$\gamma = 5$	ADMM	31.36	0.0940	31.63	200
平方	$\gamma = 5$	DCA	29.09	0.1221	19.68	126
平方	$\gamma = 5$	nmBDCA	30.10	0.1086	9.52	64
平方	$\gamma = 5$	Ours	**31.61**	**0.0914**	**5.10**	**32**
平均值	$\gamma = 3$	ADMM	29.56	0.0740	46.03	200
平均值	$\gamma = 3$	DCA	28.74	0.0813	19.19	162
平均值	$\gamma = 3$	nmBDCA	29.33	0.0759	7.10	71
平均值	$\gamma = 3$	Ours	**30.00**	**0.0714**	**5.03**	**48**
平均值	$\gamma = 5$	ADMM	25.48	0.1118	44.00	200
平均值	$\gamma = 5$	DCA	25.32	0.1138	26.71	122
平均值	$\gamma = 5$	nmBDCA	25.71	0.1019	32.32	63
平均值	$\gamma = 5$	Ours	**27.44**	**0.0953**	**6.55**	**31**

参考文献

[1]Arag´on Artacho, F.J., Vuong, P.T.：用于非光滑函数的加速凸函数之差算法。《SIAM 优化杂志》30(1)，980–1006 (2020) https://doi.org/10.1137/18M123339X https://doi.org/10.1137/18M123339X[2] Arag´on Artacho, F.J., Campoy, R., Vuong, P.T.：用于线性约束DC规划的加速DC算法。arXiv电子预印本，1908–01138 (2019) arXiv:1908.01138 [math.OC[3] An, L.T.H., Tao, P.D., Muu, L.D.：通过d.c.优化算法求解有效集上的优化问题。《运筹学快报》19(3)，117–128 (1996) https://doi.org/10.1016/0167-6377(96)00022-3[4] Tao, P.D., Souad, E.B.：求解一类非凸优化问题的算法：次梯度方法。载于：Hiriart-Urruty, J.-B.（编）《Fermat Days 85：优化数学》。北荷兰数学研究丛书，第129卷，第249–271页。North-Holland，??? (1986)。https://doi.org/10.1016/S0304-0208(08) 72402-2 . https://www.sciencedirect.com/science/article/pii/S0304020808724022[5]Le Thi, H.A., Pham Dinh, T.：DC规划与DCA：三十年的发展。《数学规划》169(1)，5–68 (2018) https://doi.org/10.1007/ s10107-018-1235-y[6] Pham Dinh, T., Le Thi, H.A.：d.c.规划的凸分析方法：理论、算法与应用。《越南数学学报》22(1)，289– 355 (1997)[7] Le Thi, H.A., Pham Dinh, T.：DC规划与DCA中的开放问题和最新进展。《全球优化杂志》88(3)，533–590 (2023)[8] Geremew, W., Nam, N.M., Semenov, A., Boginski, V., Pasiliao, E.：一种通过Nesterov光滑化技术求解多播网络设计问题的DC规划方法。《全球优化杂志》72(4)，705–729 (2018) https: // doi.org/10.1007/s10898-018-0671-9[9] Gotoh, J.-Y., Takeda, A., Tono, K.：稀疏优化问题的DC形式及算法。《数学规划》169(1)，141–176 (2018) https://doi.org/ 10.1007/s10107-017-1181-0[10]Lou, Y., Zeng, T., Osher, S., Xin, J.：一种用于图像处理的各向异性与各向同性全变分加权差分模型。SIAM影像科学杂志8(3)，1798–1823 (2015) https://doi.org/10.1137/14098435X https://doi.org/10.1137/14098435X

[11] Sankaranarayanan, P., Rengaswamy, R.：Cdinn——凸差分神经网络。《神经计算》**495**，153–168 (2022) https://doi.org/10.1016/j.neucom.2022.01.024[12] Ferreira, O.P., Santos, E.M., Souza, J.C.O.：用于DC规划的增强型缩放次梯度法。arXiv预印本 arXiv:2103.10757 (2021)[13]Bagirov, A.M., Hoseini Monjezi, N., Taheri, S.：用于最小化非光滑DC函数的增广次梯度法。《计算优化与应用》**80**(2)，411–438 (2021) https://doi.org/10.1007/s10589-021-00304-4[14] Cruz Neto, J., Oliveira, P.R., Soubeyran, A., Souza, J.：一种用于DC函数的广义近端线性化算法及其在企业最优规模问题中的应用。《运筹学年刊》**289**(2)，313–339 (2020)[15] Souza, J.C.O., Oliveira, P.R., Soubeyran, A.：凸函数之差的近端线性化算法的全局收敛性。《优化快报》**10**(7)，1529–1539 (2016)[16] Li, X., Yang, L., Ge, J., Haupt, J., Zhang, T., Zhao, T.：关于非凸稀疏学习中DC近端牛顿算法的二次收敛性。《神经信息处理系统进展》**30** (2017)[17] Oliveira, W., Tcheou, M.P.：一种DC规划惯性算法。《集值与变分分析》**27**(4)，895–919 (2019)[18] Arag´on Artacho, F.J., Fleming, R.M., Vuong, P.T.：用于光滑函数的加速DC算法。《数学规划》**169**(1)，95–118 (2018)[19]Li, Z., Lou, Y., Zeng, T.：通过DC规划去除变分乘性噪声。《科学计算杂志》**68**(3)，1200–1216 (2016)[20] Sciacchitano, F., Dong, Y., Zeng, T.：利用柯西噪声复原模糊图像的变分方法。SIAM影像科学杂志**8**(3)，1894–1922 (2015)[21] Chen, W., You, J., Chen, B., Pan, B., Li, L., Pomeroy, M., Liang, Z.：一种基于稀疏表示和字典学习、用于瑞利噪声环境下图像复原的算法。《神经计算》**286**，130–140 (2018) https://doi.org/10.1016/j.neucom.2018.01.066[22] Wu, Y., Tan, H., Li, Y., Li, F., He, H.：基于柯西分布的鲁棒张量分解及其应用。《神经计算》**223**，107–117 (2017)https://doi.org/10.1016/j.neucom.2016.10.030[23] Zhang, B., Zhu, G., Zhu, Z.：一种利用脉冲噪声图像去模糊的TV-log非凸方法。《信号处理》**174**，107631 (2020)

[24] Gu, G.,Jiang, S., Yang, J.：采用tvscad 方法处理脉冲噪声图像去模糊。《反问题》**33**（12），125008（2017）[25] Lanza, A., Morigi, S., Sgallari, F.：通过带参数选择的非凸正则化进行凸图像去噪。《数学成像与视觉杂志》**56**（2），195–220（2016）[26] Nguyen, T.B.T.：图像处理中DC规划与DCA：压缩感知、分割与复原。论文，洛林大学´（2014年12月）。https://hal.univ-lorraine.fr/tel-01751600[27] Huang, M., Lai, M.-J., Varghese, A., Xu, Z.：关于基于DC的方法在相位恢复中的应用。载于：《逼近理论国际会议》，第87–121页（2019）。施普林格[28] Li, Z., Wan, C., Tan, B., Yang, Z., Xie, S.：带SCAD惩罚的快速DC字典学习算法。《神经计算》**429**，89–100（2021）[29] Ferreira, O.P., Santos, E.M., Souza, J.C.O.：针对不可微DC分量、采用非单调线搜索的加速DC算法。arXiv预印本，arXiv:2111.01290（2021）[30] Ferreira, O.P., Rabelo, R.A.L., Ribeiro, P.H.A., Santos, E.M., Souza, J.C.O.：采用针对非凸模型的非单调增强型DCA进行图像去噪。《计算机与电气工程》**117**，109306（2024）https://doi.org/10.1016/j.compeleceng.2024.109306[31] Niu, Y.-S., Wang, Y.-J., Thi, H.A.L., Pham, D.T.：通过加速差分凸规划方法与平方和进行高阶矩投资组合优化。arXiv预印本，arXiv:1906.01509（2019）[32] Zhang, H., Niu, Y.-S.：采用幂和DC分解的线性约束多项式规划增强型DCA。《优化理论与应用杂志》**201**（2），720–759（2024）[33] Fukushima, M., Mine, H.：针对某些非凸最小化问题的广义近端点算法。《国际系统科学杂志》**12**（8），989–1000（1981）[34] Mine, H., Fukushima, M.：凸函数与连续可微函数之和的最小化方法。《优化理论与应用杂志》**33**（1），9–23（1981）[35] Pallaschke, D.：通往凸分析及其应用的简易路径。《优化》**65**，567–568（2016）[36] Evans, L.C., Garzepy, R.F.：测度论与函数的精细性质。

劳特利奇，???（2018）

[37] Clarke, F.H.：《优化与非光滑分析》。SIAM，???（1990）[38] Rockafellar, R.T.：《凸分析》第18卷。普林斯顿大学出版社，???（1970）[39] Ferreira, O.P., Santos, E.M., Souza, J.C.O.：针对非可微DC分量的加速DC算法及其非单调线搜索。计算优化与应用，1–36（2024）[40] Fan, J., Li, R.：通过非凹惩罚似然进行变量选择及其预言机性质。《美国统计学会杂志》**96**（456），1348–1360（2001）[41] Le Thi, H.A., Huynh, V., Pham Dinh, T.：含次解析数据的DC规划中DC算法的收敛性分析。《运筹学年鉴》。技术报告，LMI，鲁昂国立应用科学学院（2009）[42] Chouzenoux, E., Pesquet, J.-C., Repetti, A.：分块坐标变度量前向–后向算法。《全球优化杂志》**66**（3），457–485（2016）[43] Chai, Y.,Li, G., Qin, S.,Feng, J., Xu, C.：求解非凸资源分配问题的神经动力学优化方法。《神经计算》**512**，178–189（2022）https://doi.org/10.1016/j.neucom.2022.09.044[44] Attouch, H., Bolte, J.：关于涉及解析特征的非光滑函数近端算法的收敛性。《数学规划》**116**，5–16（2009）[45] Attouch, H., Bolte, J., Redont, P., Soubeyran, A.：用于非凸问题的近端交替最小化与投影方法：一种基于Kurdyka-Łojasiewicz不等式的方法。《运筹学数学》**35**（2），438–457（2010）[46] Attouch, H., Bolte, J., Svaiter, B.F.：半代数与温和问题中下降方法的收敛性：近端算法、前向–后向分裂及正则化高斯–赛德尔方法。《数学规划》**137**（1），91–129（2013）[47] Bolte, J., Daniilidis, A., Lewis, A.：非光滑次解析函数的 Łojasiewicz不等式及其在次梯度动力系统中的应用。《SIAM 优化杂志》**17**（4），1205–1223（2007）[48] Xie, Y., Chi, Y., Calderbank, R.：有毒噪声下的低秩矩阵恢复。载于：2013年IEEE 全球信号与信息处理大会，第622–622页（2013）。IEEE

[49] Salmon, J., Harmany, Z., Deledalle, C.-A., Willett, R.：利用非局部主成分分析降低泊松噪声。《数学成像与视觉杂志》**48**，279–294（2014）[50] Chambolle, A.：一种全变分最小化算法及其应用。《数学成像与视觉杂志》**20**，89–97（2004）[51] Mei, J.-J., Dong, Y., Huang, T.-Z., Yin, W.：通过具有收敛性保证的非凸交替方向乘子法去除柯西噪声。《科学计算杂志》**74**，743–766（2018）[52] Chambolle, A., Pock, T.：一种用于凸问题的一阶原始-对偶算法及其在成像中的应用。《数学成像与视觉杂志》**40**，120–145（2011）[53] Tao, P.D., 等：利用DC优化算法对有效集上的优化问题进行数值求解。运筹学快报**19**（3），117–128（1996）[54] Abdellatif, M., Maing´e, P.-E.：关于DC函数近似方法的收敛性。计算数学杂志**24**，475–480（2006）