# Rebuttal for the paper #2705

Anonymous

March 29, 2025


## Abstract

In this PDF, we list several major revisions for the submitted paper "Learning Cocoercive Conservative Denoisers via Helmholtz-Hodge Decomposition for Poisson Imaging Inverse Problems". The theoretical revisions are listed in Sections 1-2. All the theoretical revisions are marked in blue. The experimental revisions are listed in Sections ???-???.

In theory, we corrected the original Theorem 2.2. Besides, we fortunately find that the original characterization Theorem 2.2 is not tight enough: when $t$ is small, $\partial F$ is also Lipschitz, see the current Theorem 1.2 for details. Since Theorem 1.2 has been revised, the convergence theorem is also strengthened, see the current Theorem 2.1 for details. Now, by Theorem 1.2 and Theorem 2.1, when $\gamma = 0.25$, $0 \le t < t_0 = \frac{1}{3}$, CoCo-ADMM converges. In every old and new experiment, we uniformly set $\gamma = 0.25$ and $t = 0.33$. Thus CoCo-ADMM is guaranteed to converge.

In experiments, we .........


## 1 Better Characterization Theorem for CoCo Denoisers

We have already proved the following theorem in the original manuscript. Here we state this theorem, and omit the proof, because the proof was correct already.

**Theorem 1.1.** *Let $\mathrm{D}_\sigma \in \mathcal{C}^1[V]$. $\beta = \frac{1}{\sigma^2}$. $\mathrm{D}_\sigma$ satisfies that:*
- *$\mathrm{D}_\sigma$ is conservative;*
- *$\mathrm{D}_\sigma$ is $\gamma$-cocoercive with $\gamma \in (0, \infty)$.*

*Then, there exists a function $F : V \to \bar{\mathbb{R}}$, such that $F$ is $r$-weakly convex, where $r = \beta(1-\gamma)$, and that $\mathrm{D}_\sigma(x) \in \mathrm{Prox}_{\frac{F}{\beta}}(x), \forall x \in V$.*

The following Theorem 1.2, that was the Theorem 2.2 originally, has been corrected and strengthened. Specifically speaking, now by Theorem 1.2, for all $\gamma \in (0,1)$, $t \in [0,1)$, $\partial F$ is $L$-Lipschitz.

**Theorem 1.2.** *Let $\mathrm{D}_\sigma \in \mathcal{C}^1[V]$. $\beta = \frac{1}{\sigma^2}$. $\mathrm{D}_\sigma$ satisfies that:*
- *$\mathrm{D}_\sigma$ is conservative;*
- *$\mathrm{D}_\sigma$ is $\gamma$-cocoercive with $\gamma \in (0,1)$.*

*Let $\mathrm{D}_\sigma^t = t\,\mathrm{D}_\sigma + (1-t)\,\mathrm{I}$, $t \in [0,1)$. It holds that:*
- *there exists a $r$-weakly convex function $F : V \to \bar{\mathbb{R}}$, $r(t) = \beta\frac{t-\gamma t}{t+\gamma-\gamma t}$, such that $\mathrm{D}_\sigma^t(x) \in \mathrm{Prox}_{\frac{F}{\beta}}(x), \forall x \in V$;*
- *$\partial F$ is $L$-Lipschitz, where*

$$L(t) = \begin{cases} \beta\frac{t}{1-t} \ge r(t), & \text{if } t \ge \frac{1-2\gamma}{2-2\gamma} \text{ and } t \in [0,1) \text{ (Case 1)}; \\ r(t) = \beta\frac{t-\gamma t}{t+\gamma-\gamma t}, & \text{if } t \le \frac{1-2\gamma}{2-2\gamma} \text{ and } t \in [0,1) \text{ (Case 2)}. \end{cases} \tag{1}$$

*Proof.* Since $\mathrm{D}_\sigma$ is $\gamma$-cocoercive with $\gamma > 0$, $\mathrm{D}_\sigma^t$ is naturally $\frac{\gamma}{t+(1-t)\gamma}$-cocoercive: $\forall x, y \in V$, we have

$$\begin{aligned} \langle \mathrm{D}_\sigma(x) - \mathrm{D}_\sigma(y), x-y\rangle &\ge \gamma\|\mathrm{D}_\sigma(x) - \mathrm{D}_\sigma(y)\|^2 \\ \langle \frac{1}{t}(\mathrm{D}_\sigma^t - (1-t)\,\mathrm{I}) \circ (x) - \frac{1}{t}(\mathrm{D}_\sigma^t - (1-t)\,\mathrm{I}) \circ (y), x-y\rangle &\ge \gamma\|\tfrac{1}{t}(\mathrm{D}_\sigma^t - (1-t)\,\mathrm{I}) \circ (x) - \tfrac{1}{t}(\mathrm{D}_\sigma^t - (1-t)\,\mathrm{I}) \circ (y)\|^2 \\ t\langle \mathrm{D}_\sigma^t(x) - \mathrm{D}_\sigma^t(y), x-y\rangle - t(1-t)\|x-y\|^2 &\ge \gamma\|\mathrm{D}_\sigma^t(x) - \mathrm{D}_\sigma^t(y) - (1-t)(x-y)\|^2. \end{aligned} \tag{2}$$

Denote $a = \mathrm{D}_\sigma^t(x) - \mathrm{D}_\sigma^t(y)$, $b = x - y$ for convenience. Then we have

$$\begin{aligned} t\langle a,b\rangle - t(1-t)\|b\|^2 &\ge \gamma\|a\|^2 - 2\gamma(1-t)\langle a,b\rangle + \gamma(1-t)^2\|b\|^2, \\ (t+2\gamma(1-t))\langle a,b\rangle &\ge \gamma\|a\|^2 + (t(1-t) + \gamma(1-t)^2)\|b\|^2. \end{aligned} \tag{3}$$

Now note that

$$\langle a, b\rangle = t\langle \mathrm{D}_\sigma(x) - \mathrm{D}_\sigma(y), x-y\rangle + (1-t)\|x-y\|^2. \tag{4}$$

Since $\mathrm{D}_\sigma$ is $\gamma$-cocoercive with $\gamma > 0$, we know that $\mathrm{D}_\sigma$ is $\frac{1}{\gamma}$-Lipschitz. Therefore, by Cauchy-Schwarz inequality,

$$\langle \mathrm{D}_\sigma(x) - \mathrm{D}_\sigma(y), x-y\rangle \leq \frac{1}{\gamma}\|x-y\|^2. \tag{5}$$

That is,

$$\begin{aligned} \langle a, b\rangle &\leq \left(\tfrac{t}{\gamma} + 1 - t\right)\|b\|^2, \\ \gamma(1-t)\langle a, b\rangle &\leq t(1-t)\|b\|^2 + \gamma(1-t)^2\|b\|^2. \end{aligned} \tag{6}$$

By substituting (6) into (3), we have

$$\begin{aligned} (t+\gamma(1-t))\langle a, b\rangle &\geq \gamma\|a\|^2, \\ \langle a, b\rangle &\geq \frac{\gamma}{t+\gamma(1-t)}\|a\|^2, \end{aligned} \tag{7}$$

which means that $\mathrm{D}_\sigma^t$ is $\frac{\gamma}{t+\gamma(1-t)}$-cocoercive. Since $\mathrm{D}_\sigma$ is conservative, $\mathrm{D}_\sigma^t$ is also conservative. By Theorem 1.1, we know that there exists a $r$-weakly convex function $F : V \to \bar{\mathbb{R}}$, where

$$r = r(t) = \beta\left(1 - \frac{\gamma}{t+\gamma(1-t)}\right) = \beta\frac{t-\gamma t}{t+\gamma-\gamma t},$$

such that $\mathrm{D}_\sigma^t \in \mathrm{Prox}_{\frac{F}{\beta}}$.

Now we prove that $\partial F$ is $L$-Lipschitz. We consider two cases seperately.

**Case 1:** $t \geq \frac{1-2\gamma}{2-2\gamma}$ and $t \in [0, 1)$.

In this case, we need to prove that $\partial F$ is $L(t) = \beta\frac{t}{1-t}$ Lipschitz. Given $\forall x, y \in V$, choose arbitrary $u, v$, such that, $u = \mathrm{D}_\sigma^t(x) \in (\mathrm{I} + \frac{1}{\beta}\partial F)^{-1}(x), v = \mathrm{D}_\sigma^t(y) \in (\mathrm{I} + \frac{1}{\beta}\partial F)^{-1}(y)$, then

$$\beta(x-u) \in \partial F(u), \beta(y-v) \in \partial F(v). \tag{8}$$

Note that $a = u - v, b = x - y$. In order to prove $\partial F$ is $L(t)$-Lipschitz, we need to prove

$$\begin{aligned} \|\beta(x-u) - \beta(y-v)\|^2 &\leq L^2(t)\|u-v\|^2 = \beta^2\frac{t^2}{(1-t)^2}\|u-v\|^2, \\ \|\beta(b-a)\|^2 &\leq \beta^2\frac{t^2}{(1-t)^2}\|a\|^2, \\ \|b-a\|^2 &\leq \frac{t^2}{(1-t)^2}\|a\|^2. \end{aligned} \tag{9}$$

Since $a = u - v = \mathrm{D}_\sigma^t(x) - \mathrm{D}_\sigma^t(y) = t(\mathrm{D}_\sigma(x) - \mathrm{D}_\sigma(y)) + (1-t)(x-y)$, $b = x - y$, we have

$$a - b = t(\mathrm{D}_\sigma(x) - \mathrm{D}_\sigma(y)) - t(x-y). \tag{10}$$

Now we only need to prove

$$\begin{aligned} & t^2\|\mathrm{D}_\sigma(x) - \mathrm{D}_\sigma(y)\|^2 - 2t^2\langle \mathrm{D}_\sigma(x) - \mathrm{D}_\sigma(y), x-y\rangle + t^2\|x-y\|^2 \\ \leq\ & \frac{t^2}{(1-t)^2}\left(t^2\|\mathrm{D}_\sigma(x) - \mathrm{D}_\sigma(y)\|^2 + 2t(1-t)\langle \mathrm{D}_\sigma(x) - \mathrm{D}_\sigma(y), x-y\rangle + (1-t)^2\|x-y\|^2\right), \end{aligned} \tag{11}$$

which is equivalent to prove

$$\frac{1-2t}{2-2t}\|\mathrm{D}_\sigma(x) - \mathrm{D}_\sigma(y)\|^2 \leq \langle \mathrm{D}_\sigma(x) - \mathrm{D}_\sigma(y), x-y\rangle. \tag{12}$$

Since $t \geq \frac{1-2\gamma}{2-2\gamma}$, we have that

$$\frac{1-2t}{2-2t} \leq \frac{1-2\frac{1-2\gamma}{2-2\gamma}}{2-2\frac{1-2\gamma}{2-2\gamma}} = \frac{2-2\gamma-2(1-2\gamma)}{2(2-2\gamma)-2(1-2\gamma)} = \frac{2\gamma}{2} = \gamma. \tag{13}$$

We already have

$$\gamma\|\mathrm{D}_\sigma(x) - \mathrm{D}_\sigma(y)\|^2 \leq \langle \mathrm{D}_\sigma(x) - \mathrm{D}_\sigma(y), x-y\rangle. \tag{14}$$

Thus,

$$\frac{1-2t}{2-2t}\|\mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y)\|^2 \le \gamma\|\mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y)\|^2 \le \langle \mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y), x-y\rangle. \tag{15}$$

**Case 2 (new added): $t \le \frac{1-2\gamma}{2-2\gamma}$ and $t\in[0,1)$.**

In this case, we need to prove that $\partial F$ is $L(t)=r(t)=\beta\frac{t-t\gamma}{t+\gamma-t\gamma}$ Lipschitz. Similarly in (9), we need to prove

$$\begin{aligned} \|\beta(x-u)-\beta(y-v)\|^2 &\le L^2(t)\|u-v\|^2 = \beta^2\frac{t^2(1-\gamma)^2}{(t+\gamma-t\gamma)^2}\|u-v\|^2, \\ \|b-a\|^2 &\le \frac{t^2(1-\gamma)^2}{(t+\gamma-t\gamma)^2}\|a\|^2. \end{aligned} \tag{16}$$

Since $a=u-v=\mathrm{D}_\sigma^t(x)-\mathrm{D}_\sigma^t(y)=t(\mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y))+(1-t)(x-y)$, $b=x-y$, we have

$$a-b=t(\mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y))-t(x-y). \tag{17}$$

Now we only need to prove

$$\begin{aligned} & t^2\|\mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y)\|^2-2t^2\langle \mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y),x-y\rangle+t^2\|x-y\|^2 \\ \le\ & \frac{t^2(1-\gamma)^2}{(t+\gamma-t\gamma)^2}\left(t^2\|\mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y)\|^2+2t(1-t)\langle \mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y),x-y\rangle+(1-t)^2\|x-y\|^2\right). \end{aligned} \tag{18}$$

When $t=0$, it holds naturally. When $t\in(0,1)$, and $t\le\frac{1-2\gamma}{2-2\gamma}$, it is equivalent to prove

$$\begin{aligned} & \|\mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y)\|^2-2\langle \mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y),x-y\rangle+\|x-y\|^2 \\ \le\ & \frac{(1-\gamma)^2}{(t+\gamma-t\gamma)^2}\left(t^2\|\mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y)\|^2+2t(1-t)\langle \mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y),x-y\rangle+(1-t)^2\|x-y\|^2\right), \end{aligned} \tag{19}$$

which is equivalent to prove

$$\begin{aligned} & (t+\gamma-t\gamma)^2\left[\|\mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y)\|^2-2\langle \mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y),x-y\rangle+\|x-y\|^2\right] \\ \le\ & (1-\gamma)^2\left[t^2\|\mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y)\|^2+2t(1-t)\langle \mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y),x-y\rangle+(1-t)^2\|x-y\|^2\right], \end{aligned} \tag{20}$$

that is to prove

$$\begin{aligned} & \left[(t+\gamma-t\gamma)^2-t^2(1-\gamma)^2\right]\|\mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y)\|^2-\left[2(t+\gamma-t\gamma)^2+2t(1-t)(1-\gamma)^2\right]\langle \mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y),x-y\rangle \\ & \le\left[(1-t)^2(1-\gamma)^2-(t+\gamma-t\gamma)^2\right]\|x-y\|^2. \end{aligned} \tag{21}$$

Now we check each coefficient, and estimate each term carefully. For the coefficient of $\|\mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y)\|^2$, we have

$$(t+\gamma-t\gamma)^2-t^2(1-\gamma)^2=2\gamma(1-\gamma)t+\gamma^2>0 \tag{22}$$

For the coefficient of $\langle \mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y),x-y\rangle$, it is obviously non-positive. Besides, we have that

$$\begin{aligned} & -\left[2(t+\gamma-t\gamma)^2+2t(1-t)(1-\gamma)^2\right] \\ =\ & -2\left[(1-\gamma)^2t^2+2\gamma(1-\gamma)t+\gamma^2+(t-t^2)(1-\gamma)^2\right] \\ =\ & -(2-2\gamma^2)t-2\gamma^2. \end{aligned} \tag{23}$$

Since $\mathrm{D}_\sigma$ is $\gamma$-cocoercive, we have that

$$\left[-(2-2\gamma^2)t-2\gamma^2\right]\langle \mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y),x-y\rangle \le \left[-\gamma(2-2\gamma^2)t-2\gamma^3\right]\|\mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y)\|^2. \tag{24}$$

For the coefficient of $\|x-y\|^2$, note that when $\gamma\in(0,1)$, $t\in[0,1)$, and $t\le\frac{1-2\gamma}{2-2\gamma}$, we have

$$(1-t)^2(1-\gamma)^2-(t+\gamma-t\gamma)^2=(2\gamma-2)t-2\gamma+1\ge(2\gamma-2)\frac{1-2\gamma}{2-2\gamma}-2\gamma+1\ge 0. \tag{25}$$

Combining (22)-(25), to prove (21), we only need to prove that

$$\begin{aligned} & \left[2\gamma(1-\gamma)t+\gamma^2-\gamma(2-2\gamma^2)t-2\gamma^3\right]\|\mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y)\|^2 &\le \left[(2\gamma-2)t-2\gamma+1\right]\|x-y\|^2 \\ \iff\ & \left[(2\gamma^3-2\gamma^2)t+\gamma^2-2\gamma^3\right]\|\mathrm{D}_\sigma(x)-\mathrm{D}_\sigma(y)\|^2 &\le \left[(2\gamma-2)t-2\gamma+1\right]\|x-y\|^2. \end{aligned} \tag{26}$$

Since $\mathrm{D}_\sigma$ is $\frac{1}{\gamma}$-Lipschitz, we only need to prove that

$$\begin{aligned} & \frac{1}{\gamma^2}\left[(2\gamma^3-2\gamma^2)t+\gamma^2-2\gamma^3\right] &\le (2\gamma-2)t-2\gamma+1 \\ \iff\ & (2\gamma-2)t+1-2\gamma &\le (2\gamma-2)t-2\gamma+1. \end{aligned} \tag{27}$$

This completes the proof. □

# 2 Better Convergence Theorem for CoCo-ADMM

The restoration model is

$$\hat{u} \in \arg\min_{u\in V} F(u) + G(u; f), \quad G(u; f) = \lambda\langle \mathbf{1}, Ku - f \log Ku\rangle. \tag{28}$$

CoCo-ADMM takes the form of:

$$\begin{aligned} u^{k+1} &= \mathrm{Prox}_{\frac{G}{\beta}}(v^k - b^k), \\ v^{k+1} &= \mathrm{D}_\sigma^t(u^{k+1} + b^k), \\ b^{k+1} &= b^k + u^{k+1} - v^{k+1}, \end{aligned} \tag{29}$$

where $\mathrm{D}_\sigma^t$ is defined in Theorem 1.2, and $\beta = \frac{1}{\sigma^2}$. The PnP-ADMM algorithm in (29) with a $\gamma$-CoCo denoiser is referred to as $\gamma$-CoCo-ADMM, or CoCo-ADMM for short.

When the denoiser $\mathrm{D}_\sigma \in \mathcal{C}^1[V]$ is a CoCo denoiser satisfying the conditions in Theorem 1.2, and $F$ verifies the Kurdyka-Lojasiewicz (KL) property [1, 3], the global convergence of CoCo-ADMM in (29) can be established as follows.

Since Theorem 1.2 has been revised, the convergence theorem is also revised. In special, now by Theorem 2.1, for all $\gamma \in (0, 1)$ and $t \in [0, t_0)$, CoCo-ADMM converges. The major revision lies in part 1, where we prove that when $t < t_0$, an important value $\frac{\beta}{2} - \frac{r}{2} - \frac{L^2}{\beta}$ is positive.

**Theorem 2.1.** *Let $F : V \to \bar{\mathbb{R}}$ be a coercive weakly convex KL function in Theorem 1.2 such that $\mathrm{D}_\sigma^t = \mathrm{Prox}_{\frac{F}{\beta}}$. $G : V \to \bar{\mathbb{R}}$ is lower semi-continuous and convex. $\gamma \in (0, 1)$. $t \in [0, t_0)$, where $t_0 = t_0(\gamma)$ is the positive root of the equation*

$$(2 - 2\gamma)t^3 + \gamma t^2 + 2\gamma t - \gamma = 0. \tag{30}$$

*Then, the sequence $\{(u^k, v^k, b^k)\}$ generated by (29) converges globally to a point $(u^*, v^*, b^*)$, and that $u^* = v^*$ is a stationary point of the model (28).*

We will make use of the Lyapunov function $L_\beta$ for (29) according to [6, 9, 4]:

$$L_\beta(u, v, b) = F(v) + G(u; f) + \beta\langle b, u - v\rangle + \frac{\beta}{2}\|u - v\|^2. \tag{31}$$

We will first prove in part 1 that an important value for $L_\beta(u, v, b)$ is positive whenever $t \in (0, t_0)$, where $t_0$ is the positive root of the characteristic equation in (30). Then, we will prove in part 2 that $L_\beta$ is non-increasing with the iteration number $k$. Finally, we will prove in part 3 that CoCo-ADMM iteration in (29) converges globally to a stationary point of (28).

*Proof.* Let $h(t) = (2-2\gamma)t^3+\gamma t^2+2\gamma t-\gamma$, where $\gamma \in (0, 1)$. Note that $h$ is obviously smooth, and $h(0) = -\gamma < 0$, $h(\infty) = \infty$. Also note that when $t > 0$,

$$h'(t) = (2 - 2\gamma)t^2 + 2\gamma t + 2\gamma > 0. \tag{32}$$

Therefore, there exists a unique $t_0 > 0$, such that $h(t_0) = 0$, $h(t) > 0$ if $t > t_0$, and $h(t) < 0$ if $t \in [0, t_0)$.

**Part 1:**

We consider a characteristic value for $L_\beta(u, v, b)$ in (31): $\frac{\beta}{2} - \frac{r}{2} - \frac{L^2}{\beta}$.

By Theorem 1.2, if $t \in [0, 1)$ and $t \geq \frac{1-2\gamma}{2-2\gamma}$, we have that $r = \beta\frac{t-\gamma t}{t+\gamma-t\gamma}$ and $L = \beta\frac{t}{1-t}$. Thus we have

$$\begin{aligned} & \frac{\beta}{2} - \frac{r}{2} - \frac{L^2}{\beta} = \frac{\beta}{2} - \frac{\beta(t-\gamma t)}{2(t+\gamma-t\gamma)} - \frac{\beta t^2}{(1-t)^2} \\ = \;& \frac{\beta}{2}\left(1 - \frac{t-\gamma t}{t+\gamma-t\gamma} - \frac{2t^2}{(1-t)^2}\right) \\ = \;& \frac{\beta}{2(t+\gamma-t\gamma)(1-t)^2}\left(\gamma(1-t)^2 - 2t^2(t+\gamma-t\gamma)\right) \\ = \;& -\frac{\beta}{2(t+\gamma-t\gamma)(1-t)^2}\left((2-2\gamma)t^3 + \gamma t^2 + 2\gamma t - \gamma\right). \end{aligned} \tag{33}$$

When $0 < t < t_0$, where $t_0$ is the positive root of the characteristic equation in (30), $\frac{\beta}{2} - \frac{r}{2} - \frac{L^2}{\beta} > 0$ holds.

If $t \in [0,1)$ and $t \leq \frac{1-2\gamma}{2-2\gamma}$, we have that $L = r = \beta\frac{t-\gamma t}{t+\gamma-t\gamma}$. Note that in this case, $L = r \leq \beta\frac{t}{1-t}$. Thus,

$$
\begin{aligned}
& \frac{\beta}{2} - \frac{r}{2} - \frac{L^2}{\beta} \\
\geq \quad & \frac{\beta}{2} - \frac{\beta(t-\gamma t)}{2(t+\gamma-t\gamma)} - \frac{\beta t^2}{(1-t)^2} \\
= \quad & -\frac{\beta}{2(t+\gamma-t\gamma)(1-t)^2}\left((2-2\gamma)t^3 + \gamma t^2 + 2\gamma t - \gamma\right).
\end{aligned}
\tag{34}
$$

Therefore, we also have that when $0 < t < t_0$, $\frac{\beta}{2} - \frac{r}{2} - \frac{L^2}{\beta} > 0$ holds.

**Part 2:**

Now we prove that $L_\beta(u^k, v^k, b^k)$ is non-increasing. Before that, we show two important formulas. For $v^{k+1}$, by the first-order optimal condition, we know that

$$
\beta b^{k+1} = -\beta(v^{k+1} - u^{k+1} - b^k) \in \partial F(v^{k+1}). \tag{35}
$$

Similarly, for $u^{k+1}$, we have

$$
-\beta(u^{k+1} - v^k + b^k) \in \partial G(u^{k+1}; f). \tag{36}
$$

In order to prove that $L_\beta(u^k, v^k, b^k)$ is non-increasing, we decompose $L_\beta(u^k, v^k, b^k) - L_\beta(u^{k+1}, v^{k+1}, b^{k+1})$ into two parts:

$$
\begin{aligned}
& L_\beta(u^k, v^k, b^k) - L_\beta(u^{k+1}, v^{k+1}, b^{k+1}) \\
= \quad & L_\beta(u^k, v^k, b^k) - L_\beta(u^{k+1}, v^k, b^k) + L_\beta(u^{k+1}, v^k, b^k) - L_\beta(u^{k+1}, v^{k+1}, b^{k+1}),
\end{aligned}
\tag{37}
$$

and estimate them separately as follows. By the convexity of $G$ and the iteration form in (29), we have

$$
\begin{aligned}
& L_\beta(u^k, v^k, b^k) - L_\beta(u^{k+1}, v^k, b^k) \\
= \quad & G(u^k; f) - G(u^{k+1}; f) + \beta\langle b^k, u^k - u^{k+1}\rangle + \frac{\beta}{2}\|u^k - v^k\|^2 - \frac{\beta}{2}\|u^{k+1} - v^k\|^2 \\
= \quad & G(u^k; f) - G(u^{k+1}; f) - \langle -\beta(u^{k+1} - v^k + b^k), u^k - u^{k+1}\rangle + \beta\langle v^k - u^{k+1}, u^k - u^{k+1}\rangle \\
& + \frac{\beta}{2}\|u^k - v^k\|^2 - \frac{\beta}{2}\|u^{k+1} - v^k\|^2 \\
\geq \quad & 0 + \beta\langle v^k - u^{k+1}, u^k - u^{k+1}\rangle + \frac{\beta}{2}\langle u^k - u^{k+1}, u^k + u^{k+1} - 2v^k\rangle \\
= \quad & \beta\langle u^k - u^{k+1}, v^k - u^{k+1} - v^k + \frac{u^k + u^{k+1}}{2}\rangle \\
= \quad & \frac{\beta}{2}\|u^k - u^{k+1}\|^2.
\end{aligned}
\tag{38}
$$

By the $r$-weakly convexity of $F$, $\forall x, y \in V$, $f_y \in \partial F(y)$, we have:

$$
F(x) - F(y) \geq \langle f_y, x - y\rangle - \frac{r}{2}\|x - y\|^2. \tag{39}
$$

By the $L$-Lipschitz property of $\partial F$, $\forall x, y \in V$, $f_y \in \partial F(y)$, we have:

$$
F(x) - F(y) \leq \langle f_y, x - y\rangle + \frac{L}{2}\|x - y\|^2. \tag{40}
$$

Combining (39) and (40), we can obtain that

$$
\begin{aligned}
& L_\beta(u^{k+1},v^k,b^k) - L_\beta(u^{k+1},v^{k+1},b^{k+1}) \\
= \; & F(v^k) - F(v^{k+1}) + \beta\langle b^k, u^{k+1} - v^k\rangle - \beta\langle b^{k+1}, u^{k+1} - v^{k+1}\rangle \\
& + \frac{\beta}{2}\|u^{k+1} - v^k\|^2 - \frac{\beta}{2}\|u^{k+1} - v^{k+1}\|^2 \\
= \; & F(v^k) - F(v^{k+1}) + \beta\langle b^k, u^{k+1} - v^k\rangle - \beta\langle b^k, u^{k+1} - v^{k+1}\rangle - \beta\langle u^{k+1} - v^{k+1}, u^{k+1} - v^{k+1}\rangle \\
& + \frac{\beta}{2}\|v^k - v^{k+1}\|^2 + \beta\langle u^{k+1} - v^{k+1}, v^{k+1} - v^k\rangle \\
= \; & F(v^k) - F(v^{k+1}) + \beta\langle b^k, v^{k+1} - v^k\rangle - \beta\|u^{k+1} - v^{k+1}\|^2 \\
& + \frac{\beta}{2}\|v^k - v^{k+1}\|^2 + \beta\langle u^{k+1} - v^{k+1}, v^{k+1} - v^k\rangle \\
= \; & F(v^k) - F(v^{k+1}) - \langle \beta b^k, v^{k+1} - v^k\rangle - \beta\|u^{k+1} - v^{k+1}\|^2 + \frac{\beta}{2}\|v^k - v^{k+1}\|^2 \\
& + \beta\langle u^{k+1} - v^{k+1}, v^{k+1} - v^k\rangle \\
\geq \; & -\frac{r}{2}\|v^k - v^{k+1}\|^2 - \beta\|u^{k+1} - v^{k+1}\|^2 + \frac{\beta}{2}\|v^k - v^{k+1}\|^2 \\
= \; & \left(\frac{\beta}{2} - \frac{r}{2}\right)\|v^k - v^{k+1}\|^2 - \beta\|b^k - b^{k+1}\|^2 \\
\geq \; & \left(\frac{\beta}{2} - \frac{r}{2}\right)\|v^k - v^{k+1}\|^2 - \frac{L^2}{\beta}\|v^k - v^{k+1}\|^2 \\
= \; & \left(\frac{\beta}{2} - \frac{r}{2} - \frac{L^2}{\beta}\right)\|v^k - v^{k+1}\|^2.
\end{aligned} \tag{41}
$$

Note that the second '=' comes from the cosine rule, the first '≥' follows from the $r$-weakly convexity of $F$ as in (39), and the second '≥' results from the $L$-Lipschitz of $\partial F$ as in (40).

Combining (38) and (41), we get

$$
L_\beta(u^k,v^k,b^k) - L_\beta(u^{k+1},v^{k+1},b^{k+1}) \geq \frac{\beta}{2}\|u^k - u^{k+1}\|^2 + \left(\frac{\beta}{2} - \frac{r}{2} - \frac{L^2}{\beta}\right)\|v^k - v^{k+1}\|^2 \geq 0, \tag{42}
$$

that is, $L_\beta(u^k,v^k,b^k)$ is non-increasing.

Now we prove that $\{(u^k,v^k,b^k)\}$ is bounded. Note that $F$ and $G$ are coercive on $V$. As a result,

$$
F(u^k) + G(u^k;f) > +\infty. \tag{43}
$$

Along with $\beta b^k \in \partial F(v^k)$ and the property that $\partial F$ is $L$-Lipschitz, we arrive at

$$
\begin{aligned}
L_\beta(u^k,v^k,b^k) \; & = F(v^k) + G(u^k;f) + \beta\langle b^k, u^k - v^k\rangle + \frac{\beta}{2}\|u^k - v^k\|^2 \\
& \geq F(u^k) + G(u^k;f) - \frac{L}{2}\|u^k - v^k\|^2 + \frac{\beta}{2}\|u^k - v^k\|^2.
\end{aligned} \tag{44}
$$

Note that $L = \frac{\beta t}{1-t}$, and that $t < 0.5$. Thus, $L < \beta$. Therefore,

$$
\begin{aligned}
& F(u^k) + G(u^k;f) - \frac{L}{2}\|u^k - v^k\|^2 + \frac{\beta}{2}\|u^k - v^k\|^2 \\
= \; & F(u^k) + G(u^k;f) + \frac{1}{2}(\beta - L)\|u^k - v^k\|^2 \geq -\infty.
\end{aligned} \tag{45}
$$

Since $F(u) + G(u;f)$ is coercive on $V$, $u^k, v^k, b^k$ are bounded.

**Part 3:**

Define $q^{k+1}$ as follows:

$$
q^{k+1} = [\beta(b^{k+1} - b^k + v^k - v^{k+1}), \beta(v^{k+1} - u^{k+1}), \beta(b^{k+1} - b^k)]. \tag{46}
$$

Define $\partial L_\beta(u,v,b) = [\partial_u L_\beta, \partial_v L_\beta, \partial_b L_\beta]$. By the formulas (35)-(36), we know that

$$
q^{k+1} \in \partial L_\beta(u^{k+1},v^{k+1},b^{k+1}). \tag{47}
$$

Note that

$$
\|\beta(b^{k+1} - b^k + v^k - v^{k+1})\| \leq \beta\|b^k - b^{k+1}\| + \beta\|v^k - v^{k+1}\| \leq L\|v^k - v^{k+1}\| + \beta\|v^k - v^{k+1}\|, \tag{48}
$$

and that

$$\|\beta(v^{k+1} - u^{k+1})\| = \beta\|b^k - b^{k+1}\| \leq L\|v^k - v^{k+1}\|, \tag{49}$$

we arrive at

$$\|q^{k+1}\| \leq C\|v^k - v^{k+1}\|, \tag{50}$$

where

$$C = 3L + \beta. \tag{51}$$

Now we can finally prove Theorem 2.1. By Part 2, $\{(u^k, v^k, b^k)\}$ is bounded. So there is a sub-sequence $\{(u^{n_k}, v^{n_k}, b^{n_k})\}$ such that $(u^{n_k}, v^{n_k}, b^{n_k}) \to (u^*, v^*, b^*)$, when $n \to +\infty$. Since $L_\beta(u^k, v^k, b^k)$ is lower bounded and non-increasing, we have that $\|u^k - u^{k+1}\|, \|v^k - v^{k+1}\| \to 0$ as $k \to +\infty$. Besides, since $q^k \in \partial L_\beta(u^k, v^k, b^k)$ and $\|q^k\| \leq C\|v^k - v^{k+1}\|$, we know that $\|q^k\| \to 0$, $\|q^{n_k}\| \to 0$. Thus, $0 \in \partial L_\beta(u^*, v^*, b^*)$, and $(u^*, v^*, b^*)$ is a stationary point of $L_\beta$.

Since $F, G$ is KL, we conclude that $L_\beta$ is also KL. Then, by the proof of Theorem 2.9 in [2], $\{(u^{n_k}, v^{n_k}, b^{n_k})\}$ converges globally to $(u^*, v^*, b^*)$.

Since $(u^*, v^*, b^*)$ is a stationary point of $L_\beta$, and as a result, $q^* = 0$, that is

$$q^* = [0, \beta(v^* - u^*), 0] = 0. \tag{52}$$

Therefore,$u^* = v^*$. By CoCo-ADMM iteration in (29), we know that

$$\begin{aligned} u^* &= \mathrm{Prox}_{\frac{G}{\beta}}(u^* - b^*), \\ u^* &= \mathrm{D}_\sigma^t(u^* + b^*) \in \mathrm{Prox}_{\frac{G}{\beta}}(u^* + b^*). \end{aligned} \tag{53}$$

Equivalently,

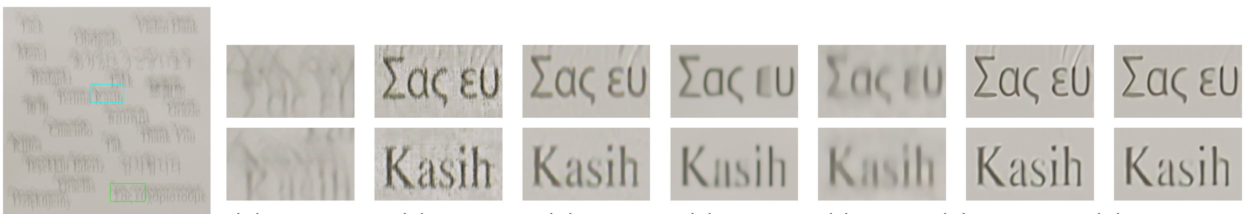
$$-\beta b^* \in \partial G(u^*; f), \beta b^* \in \partial F(u^*), \tag{54}$$

$$0 \in \partial F(u^*) + \partial G(u^*; f). \tag{55}$$

Therefore, $u^*$ is a stationary point of (28). □

# 3 Real Deblur Results

In this section, we evaluate the proposed two algorithms: CoCo-ADMM and CoCo-PEGD on a real-blur dataset containing 18 images from [5]. We compare six methods: Deblur-INR [11], Lv's method [7], DPIR [12], RMMO [8], CoCo-ADMM (Ours) and CoCo-PEGD (Ours). The first method Deblur-INR is a self-supervised deblurring method, which can estimate both clean images and blur kernels. The other four methods need blur kernels as one of the inputs. we utilize a blind deblurring method [7] to estimate approximate kernels. Visual comparisons are shown in Fig 1. We also provide quantitative comparisons on an indicator, CLIP [10] without real labels. Results are illustrated in Table 1.



(a) Blurred (b) Blurred (c) INR (d) Lv's (e) DPIR (f) RMMO (g) CoCo-ADMM (h) CoCo-PEGD

Figure 1: Blind deblurred results on real-blur datasets with five methods.

| | INR | Lv's | DPIR | RMMO | CoCo-ADMM | CoCo-PEGD |
|---|---|---|---|---|---|---|
| Average | 0.2990 | 0.3444 | 0.3629 | 0.2664 | 0.5461 | **0.5474** |

Table 1: Results comparison on binary dataset

# References


[1] Hédy Attouch, Jérôme Bolte, Patrick Redont, and Antoine Soubeyran. Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the kurdyka-łojasiewicz inequality. *Mathematics of operations research*, 35(2):438–457, 2010.

[2] Hedy Attouch, Jérôme Bolte, and Benar Fux Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized gauss–seidel methods. *Mathematical Programming*, 137(1):91–129, 2013.

[3] Pierre Frankel, Guillaume Garrigos, and Juan Peypouquet. Splitting methods with variable metric for kurdyka–łojasiewicz functions and general convergence rates. *Journal of Optimization Theory and Applications*, 165:874–900, 2015.

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

[5] Wei-Sheng Lai, Jia-Bin Huang, Zhe Hu, Narendra Ahuja, and Ming-Hsuan Yang. A comparative study for single image blind deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1701–1709, 2016.

[6] Guoyin Li and Ting Kei Pong. Douglas–rachford splitting for nonconvex optimization with application to nonconvex feasibility problems. *Mathematical programming*, 159:371–401, 2016.

[7] Jun Liu, Ming Yan, and Tieyong Zeng. Surface-aware blind image deblurring. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(3):1041–1055, 2019.

[8] Matthieu Terris, Audrey Repetti, Jean-Christophe Pesquet, and Yves Wiaux. Building firmly nonexpansive convolutional neural networks. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8658–8662. IEEE, 2020.

[9] Andreas Themelis and Panagiotis Patrinos. Douglas–rachford splitting and admm for nonconvex optimization: Tight convergence results. *SIAM Journal on Optimization*, 30(1):149–181, 2020.

[10] Jianyi Wang, Kelvin CK Chan, and Chen Change Loy. Exploring clip for assessing the look and feel of images. In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, pages 2555–2563, 2023.

[11] Tianjing Zhang, Yuhui Quan, and Hui Ji. Cross-scale self-supervised blind image deblurring via implicit neural representation. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024.

[12] Yuanzhi Zhu, Kai Zhang, Jingyun Liang, Jiezhang Cao, Bihan Wen, Radu Timofte, and Luc Van Gool. Denoising diffusion models for plug-and-play image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1219–1229, 2023.