Faculty of Engineering, Department of Computer Science

Erasmus+ International PhD Summer School 2025

Mathematics and Machine Learning for Image Analysis

University of Bologna

10 June 2025

UCL

This is joint work with

Tatiana Bubba
Università degli Studi di Ferrara

Luca Ratti
Università di Bologna

Danilo Riccio
Queen Mary University of London

Acknowledgements:

Isaac Newton Institute for Mathematical Sciences

Queen Mary University of London
Digital Environment Research Institute

The Alan Turing Institute

EPSRC

INδAM

Learning optimal sampling strategies for MRI

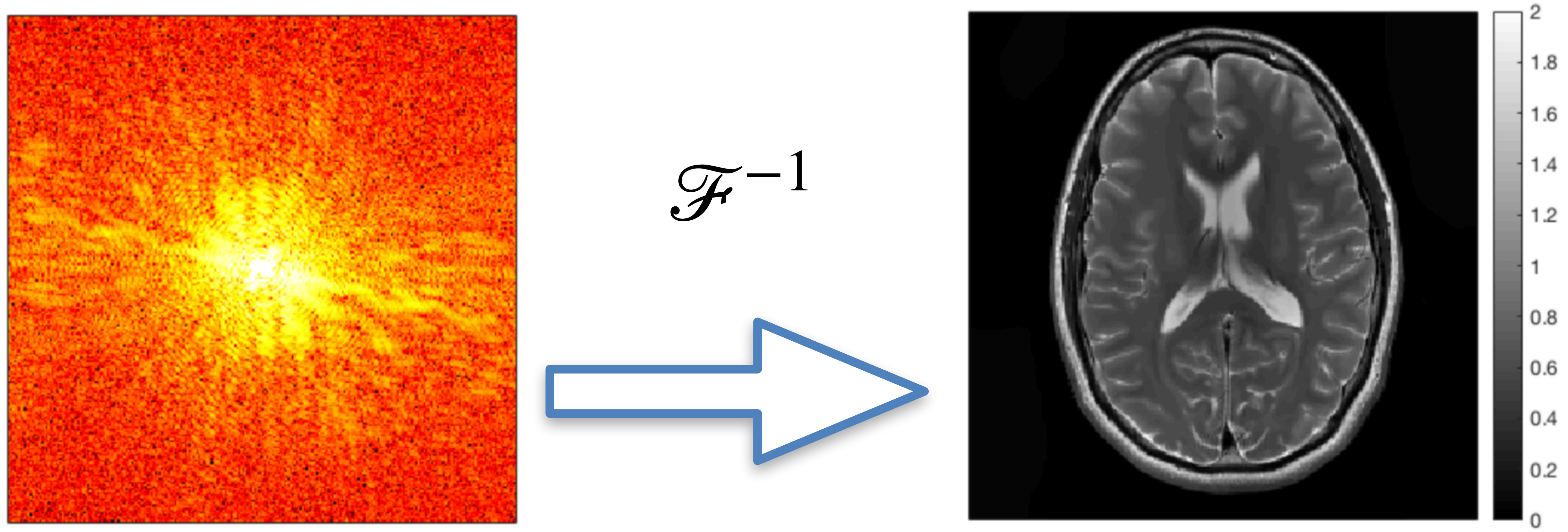
- Error estimates in inverse problems
- Computing source condition elements
- Learning variational regularisations with optimal error estimates
- Learning the sampling pattern in MRI
- Conclusions & outlook

Motivation

Magnetic resonance imaging

© Wikimedia commons

Simplified image reconstruction process



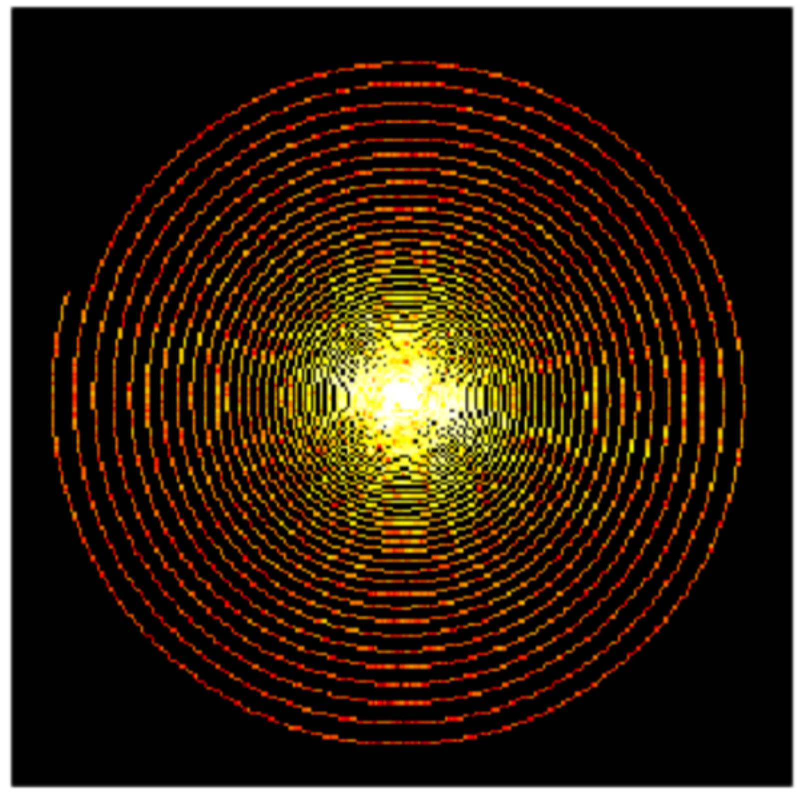
$$(\mathscr{F}u)_{pq} = \frac{1}{\sqrt{n_x n_y}} \sum_{l=0}^{n_y-1} \sum_{j=0}^{n_x-1} u_{lj}\, e^{-i\frac{2\pi pl}{n_y}} e^{-i\frac{2\pi qj}{n_x}}$$

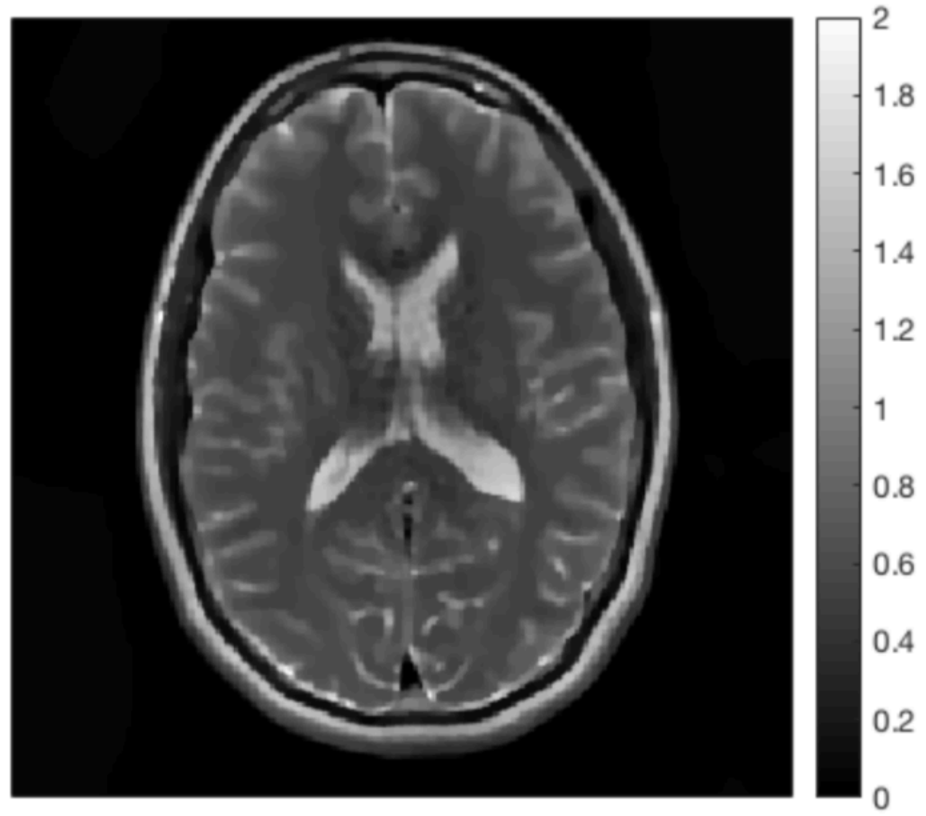
Motivation

Magnetic resonance imaging

© Wikimedia commons

Simplified image reconstruction process





Goal: speed up acquisition process

$$(\mathscr{F}u)_{pq} = \frac{1}{\sqrt{n_x n_y}} \sum_{l=0}^{n_y-1} \sum_{j=0}^{n_x-1} u_{lj}\, e^{-i\frac{2\pi pl}{n_y}} e^{-i\frac{2\pi qj}{n_x}}$$

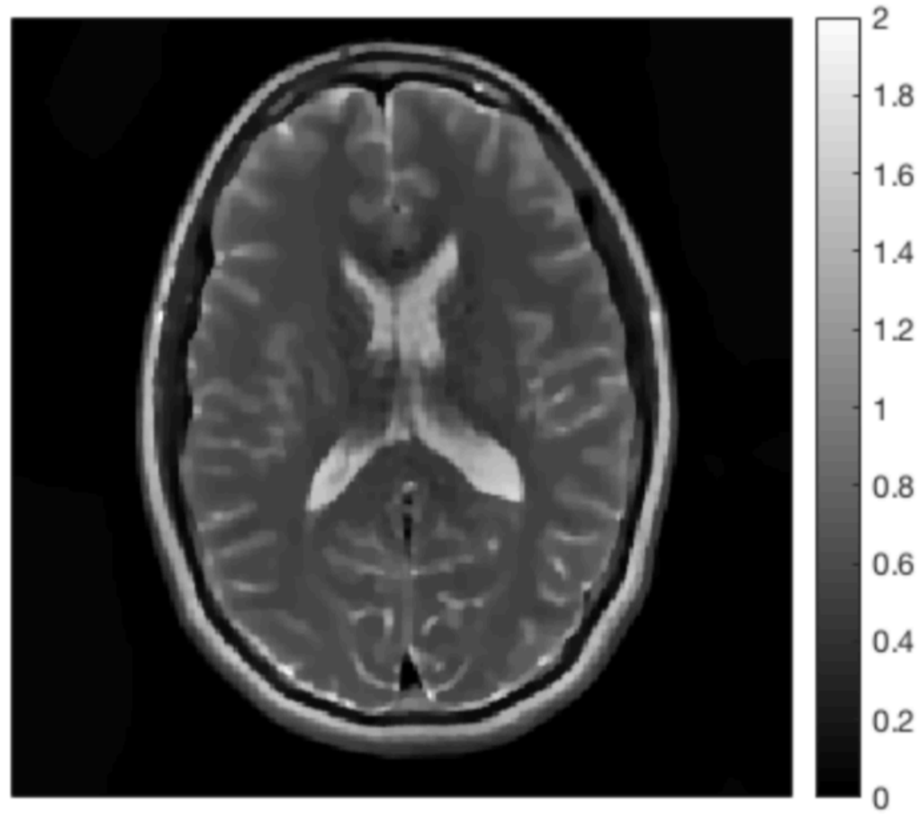
Motivation

Magnetic resonance imaging

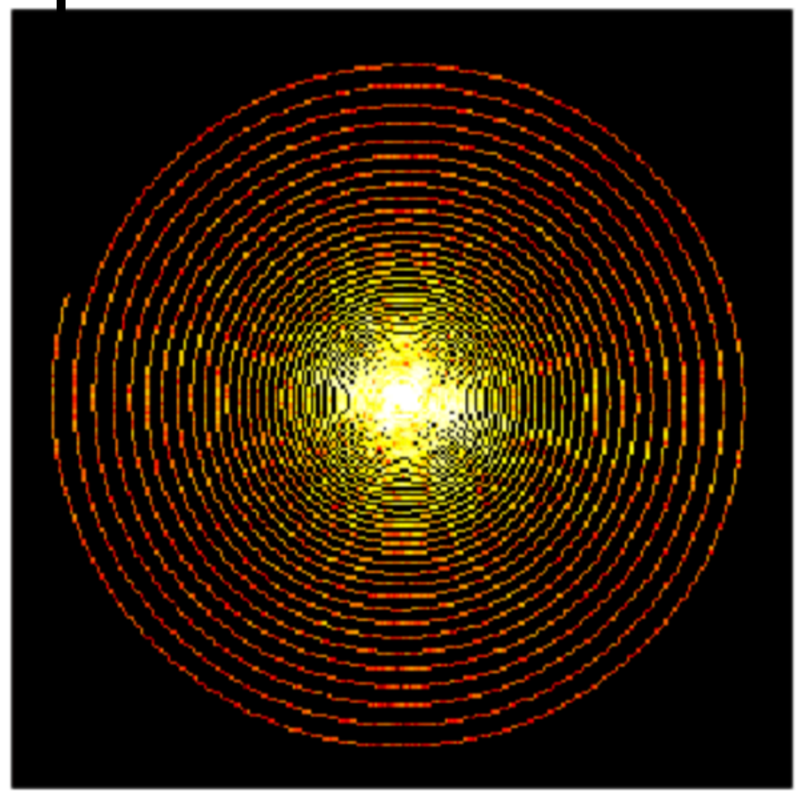
© Wikimedia commons

Simplified image formation process

Sampling operator



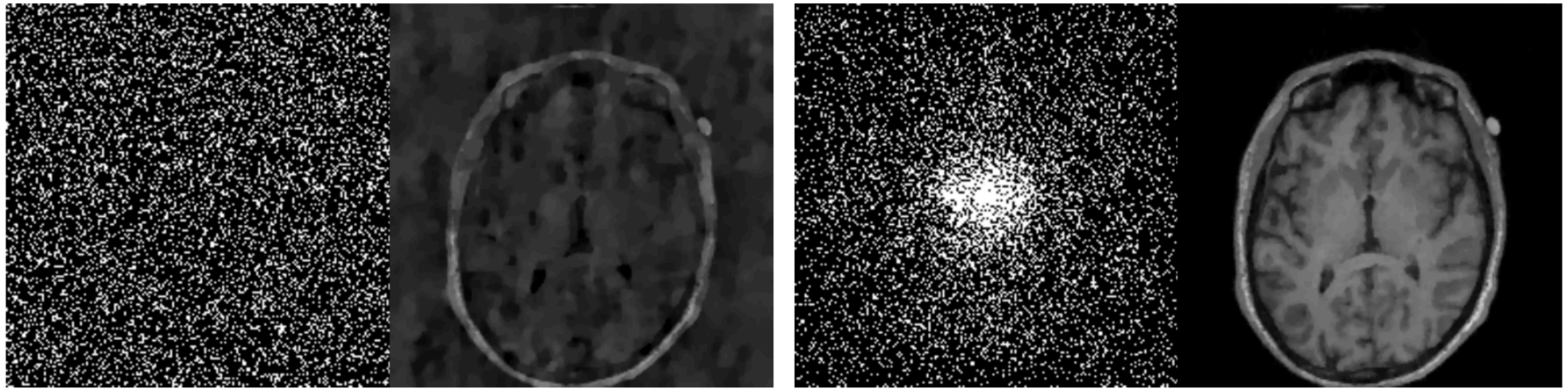
$S\mathcal{F}$



Goal: speed up acquisition process

$$(\mathcal{F}u)_{pq} = \frac{1}{\sqrt{n_x n_y}} \sum_{l=0}^{n_y-1} \sum_{j=0}^{n_x-1} u_{lj}\, e^{-i\frac{2\pi pl}{n_y}} e^{-i\frac{2\pi qj}{n_x}}$$

Motivation

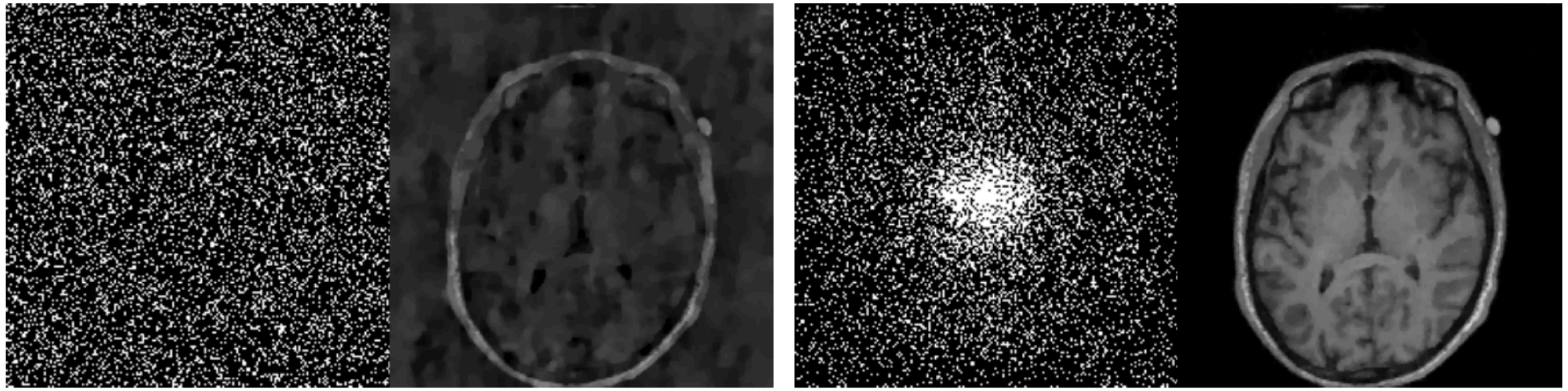


Uniform random   Reconstruction   Learned   Reconstruction

from*

*Sherry, F., MB, De los Reyes, J. C., Graves, M. J., Maierhofer, G., Williams, G., Schönlieb, C.-B. & Ehrhardt, M. J. (2020). Learning the sampling pattern for MRI. *IEEE Transactions on Medical Imaging*, *39*(12), 4310-4321.

Motivation



Uniform random | Reconstruction | Learned | Reconstruction

from*

Our approach back then:

$$\min_{S,\alpha} \|R_{S,\alpha}(f^\delta) - u^\dagger\|^2 + \mathscr{R}(S)$$

subject to

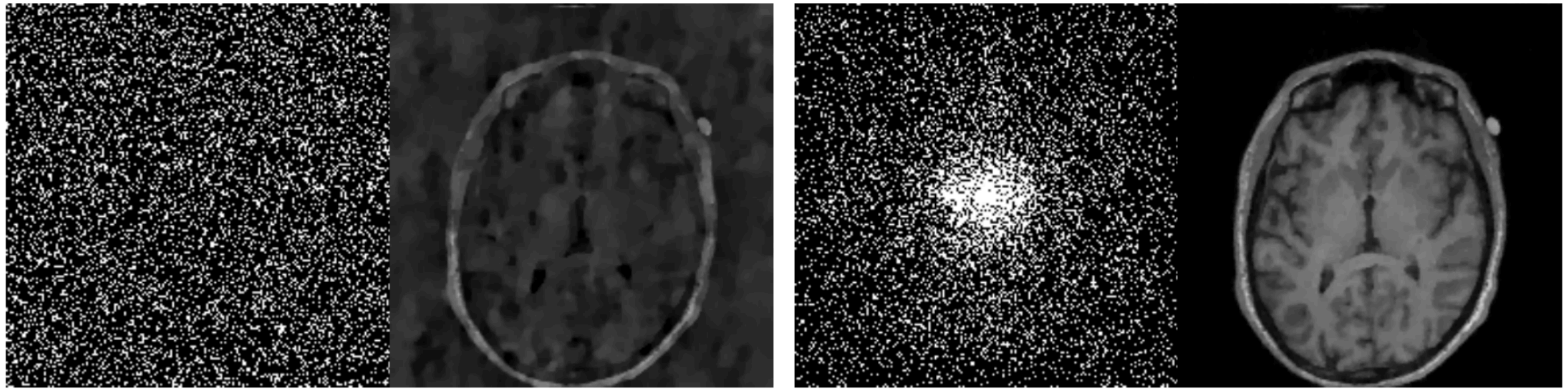
$$R_{S,\alpha}(f^\delta) \in \arg\min_u \left\{ \frac{1}{2}\|S\mathscr{F}u - f^\delta\|^2 + \alpha J(u) \right\}$$

Complex bilevel optimisation approach

Today's talk:
simple convex minimisation problem(s)

*Sherry, F., MB, De los Reyes, J. C., Graves, M. J., Maierhofer, G., Williams, G., Schönlieb, C.-B. & Ehrhardt, M. J. (2020). Learning the sampling pattern for MRI. *IEEE Transactions on Medical Imaging*, *39*(12), 4310-4321.

Motivation



Uniform random Reconstruction Learned Reconstruction

from*

Many other relevant works:

Gözcü et al. 2018, 2019 , Sanchez et al. 2019, Weiss et al. 2019, Bahadir et al. 2019, Jin, Unser, Yi 2019, Gossard, Gournay, Weiss 2022, Bakker et al. 2022, etc.

*Sherry, F., MB, De los Reyes, J. C., Graves, M. J., Maierhofer, G., Williams, G., Schönlieb, C.-B. & Ehrhardt, M. J. (2020). Learning the sampling pattern for MRI. *IEEE Transactions on Medical Imaging*, *39*(12), 4310-4321.

Error estimates in inverse problems

Error estimates in inverse problems

If we approximate the solution of an inverse problem $Ku^\dagger = f$ with regularisations R_α, we ideally want convergent regularisations

We would like to establish

$$D\left(u^\dagger, R_\alpha(f^\delta)\right) \leq C\,\delta$$

such that

$$\limsup_{\delta \to 0} \left\{ D\left(u^\dagger, R_\alpha(f^\delta)\right) \,\middle|\, \text{ for } f = Ku^\dagger \text{ and } f^\delta \text{ with } \|f - f^\delta\| \leq \delta \right\} = 0$$

for some distance measure D (not necessarily a metric)

Engl, H. W., Hanke, M., & Neubauer, A. (1996). *Regularization of inverse problems* (Vol. 375). Springer Science & Business Media.
Scherzer, O., Grasmair, M., Grossauer, H., Haltmeier, M., & Lenzen, F. (2009). Variational methods in imaging.
MB, & Burger, M. (2018). Modern regularization methods for inverse problems. *Acta Numerica*, *27*, 1-111.

Error estimates in inverse problems

If we approximate the solution of an inverse problem $Ku^\dagger = f$ with regularisations R_α, we ideally want convergent regularisations

We would like to establish

$$D\left(u^\dagger, R_\alpha(f^\delta)\right) \leq C\,\delta$$

Making this small subject to constraints is a surrogate to the previous upper level problem

such that

$$\limsup_{\delta \to 0} \left\{ D\left(u^\dagger, R_\alpha(f^\delta)\right) \;\middle|\; \text{for } f = Ku^\dagger \text{ and } f^\delta \text{ with } \|f - f^\delta\| \leq \delta \right\} = 0$$

for some distance measure D (not necessarily a metric)

Engl, H. W., Hanke, M., & Neubauer, A. (1996). *Regularization of inverse problems* (Vol. 375). Springer Science & Business Media.
Scherzer, O., Grasmair, M., Grossauer, H., Haltmeier, M., & Lenzen, F. (2009). Variational methods in imaging.
MB, & Burger, M. (2018). Modern regularization methods for inverse problems. *Acta Numerica*, *27*, 1-111.

Error estimates in inverse problems

We would like to establish

$$D\left(u^{\dagger}, R_{\alpha}(f^{\delta})\right) \leq C\,\delta$$

Example: variational regularisation

$$R_{\alpha}: f^{\delta} \mapsto u_{\alpha} \in \arg\min_{u \in X} \left\{ \frac{1}{2}\|Ku - f^{\delta}\|_{Y}^{2} + \alpha J(u) \right\}$$

Assume source (or range) condition:

$$\exists\, v \in Y: \qquad K^{*}v \in \partial J(u^{\dagger}) = \{p \mid J(v) - J(u^{\dagger}) - \langle p, v - u^{\dagger}\rangle \geq 0, \forall v\}$$

Error estimates in inverse problems

We would like to establish

$$D\left(u^{\dagger}, R_\alpha(f^\delta)\right) \leq C\,\delta$$

Example: variational regularisation

$$R_\alpha : f^\delta \mapsto u_\alpha \in \arg\min_{u \in X} \left\{ \frac{1}{2}\|Ku - f^\delta\|_Y^2 + \alpha J(u) \right\}$$

Assume source (or range) condition:

$$\exists\, v \in Y: \qquad K^* v = \nabla J(u^{\dagger})$$

Error estimates in inverse problems

We would like to establish

$$D\left(u^\dagger, R_\alpha(f^\delta)\right) \leq C\,\delta$$

Example: variational regularisation

$$R_\alpha : f^\delta \mapsto u_\alpha \in \arg\min_{u \in X} \left\{ \frac{1}{2}\|Ku - f^\delta\|_Y^2 + \alpha J(u) \right\}$$

Assume source (or range) condition:

$$\exists\, v \in Y: \qquad K^* v \in \partial J(u^\dagger)$$

Error estimates in inverse problems

We would like to establish

$$D\left(u^\dagger, R_\alpha(f^\delta)\right) \leq C\,\delta$$

Example: variational regularisation

$$R_\alpha : f^\delta \mapsto u_\alpha \in \arg\min_{u\in X} \left\{ \frac{1}{2}\|Ku - f^\delta\|_Y^2 + \alpha J(u) \right\}$$

Assume source (or range) condition:

$$\exists\, v \in Y: \qquad K^* v \in \partial J(u^\dagger)$$

(equivalent to $u^\dagger \in \arg\min_{u\in X} \left\{ \frac{1}{2}\|Ku - g_\alpha\|_Y^2 + \alpha J(u) \right\}$ with $g_\alpha = \alpha v + K u^\dagger = \alpha v + f$)

Error estimates in inverse problems

We would like to establish

$$D\left(u^\dagger, R_\alpha(f^\delta)\right) \leq C\,\delta$$

Example: variational regularisation

$$R_\alpha : f^\delta \mapsto u_\alpha \in \arg\min_{u\in X} \left\{ \frac{1}{2}\|Ku - f^\delta\|_Y^2 + \alpha J(u) \right\}$$

Assume source (or range) condition:

$$\exists\, v \in Y: \qquad K^* v \in \partial J(u^\dagger)$$

Then one can prove

$$D\left(R_{\alpha(\delta)}(f^\delta), u^\dagger\right) \leq \|v\|_Y\,\delta \qquad \text{for } \alpha(\delta) = \delta/\|v\|_Y$$

where D is a suitable (symmetrised) Bregman distance/divergence w.r.t J

Motivation

This presentation in a nutshell

Example:

$$D\left(R_{\alpha(\delta)}(f^{\delta}), u^{\dagger}\right) \leq \|v\|_Y \, \delta$$

we want to study this quantity

$u^{\dagger}$

$$K = I$$

$$J(u) = \mathrm{TV}(u) \quad *$$

*To be defined later, if you haven't seen this before

v

Motivation

This presentation in a nutshell

Example:

$$D\left(R_{\alpha(\delta)}(f^\delta), u^\dagger\right) \le \|v\|_Y \, \delta$$

we want to study this quantity

$= \alpha$

$+$

g_α $\qquad v \qquad$ $u^\dagger$

Motivation

This presentation in a nutshell

Example:

$$D\left(R_{\alpha(\delta)}(f^\delta), u^\dagger\right) \leq \|v\|_Y \,\delta$$

R_α () $=$ 

g_α $\qquad$ $u^\dagger$

Ideally, we want

$v =$ 

to have small norm $\|v\|_Y$

How do we compute v in order to estimate (or even control) $\|v\|_Y$?

Computing source condition elements

Computing source condition elements

Approach 1: we can modify the source condition

$$K^*v \in \partial J(u^\dagger)$$

Computing source condition elements

Approach 1: we can modify the source condition to

$$u^\dagger + K^* v \in \partial\left(\frac{1}{2}\|\cdot\|^2 + J\right)(u^\dagger)$$

Now we use the Fenchel-Young equality:

$$p \in \partial F(u) \quad \Longleftrightarrow \quad F(u) + F^\star(p) = \langle u, p\rangle$$

where $F^\star$ denotes the convex conjugate of F, i.e.

$$F^\star(p) := \sup_u \langle u, p\rangle - F(u)$$

Werner Fenchel, *Convex cones, sets, and functions*. Princeton University, 1953
Theorem 23.5, Ralph Tyrell Rockafellar, Convex analysis, Princeton university press, 1970

Computing source condition elements

Now we use the Fenchel-Young equality:

$$p \in \partial F(u) \qquad \Longleftrightarrow \qquad F(u) + F^\star(p) = \langle u, p \rangle$$

where $F^\star$ denotes the convex conjugate of F, i.e.

$$F^\star(p) := \sup_u \langle u, p \rangle - F(u)$$

Proof: $p \in \partial F(u) \Longleftrightarrow \quad p \in \{q \,|\, F(v) \geq F(u) + \langle q, v-u \rangle, \forall v\}$

$$\Longleftrightarrow \quad \langle p, u \rangle - F(u) \geq \langle p, v \rangle - F(v) \qquad \forall v$$

$$\Longleftrightarrow \quad \langle p, u \rangle - F(u) = F^\star(p)$$

$$\Longleftrightarrow \quad F(u) + F^\star(p) = \langle u, p \rangle$$

Werner Fenchel, *Convex cones, sets, and functions*. Princeton University, 1953
Theorem 23.5, Ralph Tyrell Rockafellar, Convex analysis, Princeton university press, 1970

Computing source condition elements

Approach 1: we can modify the source condition to

$$u^{\dagger} + K^{*}v \in \partial\left(\frac{1}{2}\|\cdot\|^{2} + J\right)(u^{\dagger})$$

Now we use the Fenchel-Young equality:

$$p \in \partial F(u) \qquad \Longleftrightarrow \qquad F(u) + F^{\star}(p) = \langle u, p\rangle$$

This implies

$$\left(\frac{1}{2}\|\cdot\|^{2} + J\right)(u^{\dagger}) + \left(\frac{1}{2}\|\cdot\|^{2} + J\right)^{\star}\left(u^{\dagger} + K^{*}v\right) = \langle u^{\dagger}, u^{\dagger} + K^{*}v\rangle$$

Werner Fenchel, *Convex cones, sets, and functions*. Princeton University, 1953
Theorem 23.5, Ralph Tyrell Rockafellar, Convex analysis, Princeton university press, 1970

Computing source condition elements

Instead of enforcing a strict equality, we can define

$$G_J(v) := \left(\frac{1}{2}\|\cdot\|^2 + J\right)(u^\dagger) + \left(\frac{1}{2}\|\cdot\|^2 + J\right)^\star \left(u^\dagger + K^*v\right) - \langle u^\dagger, u^\dagger + K^*v\rangle$$

MB, Tatiana A. Bubba, Luca Ratti, and Danilo Riccio. "Trust your source: quantifying source condition elements for variational regularisation methods." IMA Journal of Applied Mathematics (2024): hxae008.

Computing source condition elements

Instead of enforcing a strict equality, we can define

$$G_J(v) := \left(\frac{1}{2}\|\cdot\|^2 + J\right)(u^\dagger) + \left(\frac{1}{2}\|\cdot\|^2 + J\right)^{\star}(u^\dagger + K^*v) \;-\; \langle u^\dagger, u^\dagger + K^*v\rangle$$

Proposition: the gradient ∇G_J of G_J exists and reads

$$\boxed{\nabla G_J(v) = K\,\mathrm{prox}_J(u^\dagger + K^*v) - Ku^\dagger.}$$

Here prox denotes the proximal map, i.e.

$$\mathrm{prox}_F : Z \to Z, \qquad \mathrm{prox}_F(z) := \arg\min_{u \in Z} \left\{ \frac{1}{2}\|u - z\|_Z^2 + F(u) \right\}$$

MB, Tatiana A. Bubba, Luca Ratti, and Danilo Riccio. "Trust your source: quantifying source condition elements for variational regularisation methods." IMA Journal of Applied Mathematics (2024): hxae008.

Computing source condition elements

Instead of enforcing a strict equality, we can define

$$G_J(v) := \left(\frac{1}{2}\|\cdot\|^2 + J\right)(u^\dagger) + \left(\frac{1}{2}\|\cdot\|^2 + J\right)^\star \left(u^\dagger + K^*v\right) \; - \; \langle u^\dagger, u^\dagger + K^*v\rangle$$

Because of the Fréchet-differentiability of G_J, we can solve

$$\hat{v} = \arg\min_v G_J(v)$$

for instance via gradient descent, i.e

$$v^{k+1} = v^k - \tau K\left(\mathrm{prox}_J\left(u^\dagger + K^*v^k\right) - u^\dagger\right)$$

which is globally convergent for $\tau \leq 1/\|K\|^2$

MB, Tatiana A. Bubba, Luca Ratti, and Danilo Riccio. "Trust your source: quantifying source condition elements for variational regularisation methods." IMA Journal of Applied Mathematics (2024): hxae008.

Extension to more general functions and range conditions

We can also consider composite functionals

$$J(u) = H(Au + b)$$

Example: $J = \mathrm{TV}\colon \mathbb{R}^{n_y \times n_x} \to \mathbb{R}$ with

$$\mathrm{TV}(u) = \sum_{i=1}^{n_y-1} \sum_{j=1}^{n_x-1} \sqrt{\left|u_{(i+1)j} - u_{ij}\right|^2 + \left|u_{i(j+1)} - u_{ij}\right|^2}$$

We obtain this for $b = 0,\ A\colon \mathbb{R}^{n_y \times n_x} \to \mathbb{R}^{(n_y-1)\times(n_x-1)\times 2}$ with

$$(Au)_{ijp} = \begin{cases} u_{(i+1)j} - u_{ij} & p = 1 \\ u_{i(j+1)} - u_{ij} & p = 2 \end{cases},$$

and $H\colon \mathbb{R}^{(n_y-1)\times(n_x-1)\times 2} \to \mathbb{R}$ with $H(q) = \sum_{i=1}^{n_y-1} \sum_{j=1}^{n_x-1} \sqrt{\left|q_{ij1}\right|^2 + \left|q_{ij2}\right|^2}$.

MB, Tatiana A. Bubba, Luca Ratti, and Danilo Riccio. "Trust your source: quantifying source condition elements for variational regularisation methods." IMA Journal of Applied Mathematics (2024): hxae008.

Extension to more general functions and range conditions

We can also consider composite functionals

$$J(u) = H(Au + b)$$

Consider the range condition

$$u^\dagger \in \arg\min_{u \in X} \left\{ \frac{1}{2}\|Ku - g_\alpha\|^2 + \alpha J(u) \right\} ,$$

respectively it's optimality condition

$$K^*(Ku^\dagger - g_\alpha) + \alpha A^* q^\dagger = 0$$

for $q^\dagger \in \partial H(Au^\dagger + b)$.

MB, Tatiana A. Bubba, Luca Ratti, and Danilo Riccio. "Trust your source: quantifying source condition elements for variational regularisation methods." IMA Journal of Applied Mathematics (2024): hxae008.

Extension to more general functions and range conditions

Consider the range condition

$$K^*(Ku^\dagger - g_\alpha) + \alpha A^* q^\dagger = 0$$

for $q^\dagger \in \partial H(Au^\dagger + b)$.

Multiplying by $1/\alpha$ and using $g_\alpha = \alpha v + Ku^\dagger$ then yields

$$K^* v = A^* q^\dagger$$

$$q^\dagger \in \partial H(Au^\dagger + b)$$

MB, Tatiana A. Bubba, Luca Ratti, and Danilo Riccio. "Trust your source: quantifying source condition elements for variational regularisation methods." IMA Journal of Applied Mathematics (2024): hxae008.

Extension to more general functions and range conditions

Multiplying by $1/\alpha$ and using $g_\alpha = \alpha v + Ku^\dagger$ then yields

$$K^* v = A^* q^\dagger$$

$$q^\dagger \in \partial H(Au^\dagger + b)$$

Similar to before, we relax these problems by introducing $G_H\colon\ Y \to \mathbb{R}$ with

$$G_H(q) = \left(\frac{1}{2}\|\cdot\|^2 + H\right)(Au^\dagger + b) + \left(\frac{1}{2}\|\cdot\|^2 + H\right)^\star \left(q + Au^\dagger + b\right) - \langle Au^\dagger + b, Au^\dagger + b + q\rangle$$

and minimise

$$E_H(v, q^\dagger) = \frac{1}{2}\|K^* v - A^* q^\dagger\|^2 + G_H(q^\dagger)$$

MB, Tatiana A. Bubba, Luca Ratti, and Danilo Riccio. "Trust your source: quantifying source condition elements for variational regularisation methods." IMA Journal of Applied Mathematics (2024): hxae008.

Computing source condition elements

Approach 2: we can formulate the J-minimising solution of $Ku = f$, i.e.

$$\min_u J(u) \qquad \text{subject to} \qquad Ku = f$$

as a primal-dual problem

$$\min_u \sup_v J(u) + \langle v, f - Ku \rangle$$

with saddle point $(u^\dagger, v^\dagger)$ and optimality conditions

$$K^* v^\dagger \in \partial J(u^\dagger) \qquad \text{source condition}$$

$$Ku^\dagger = f$$

Computing source condition elements

Approach 2: we can formulate the J-minimising solution of $Ku = f$, i.e.

$$\min_u J(u) + H(f - Ku)$$

as a primal-dual problem

$$\min_u \sup_v J(u) + \langle v, f - Ku \rangle - H^\star(v)$$

with saddle point $(u^\dagger, v^\dagger)$ and optimality conditions

$$K^* v^\dagger \in \partial J(u^\dagger) \qquad \text{source condition}$$

$$f - Ku^\dagger \in \partial H^\star(v^\dagger)$$

Computing source condition elements

Approach 2: we can formulate the J-minimising solution of $Ku = f$, i.e.

$$\min_u J(u) + \alpha \| f - Ku \|_Y$$

as a primal-dual problem

$$\min_u \sup_v J(u) + \langle v, f - Ku \rangle - \chi_{\|\cdot\|_Y \leq \alpha}(v)$$

with saddle point $(u^\dagger, v^\dagger)$ and optimality conditions

$$K^* v^\dagger \in \partial J(u^\dagger) \qquad \text{source condition}$$

$$f - Ku^\dagger \in \partial \chi_{\|\cdot\|_Y \leq \alpha}(v^\dagger)$$

Computing source condition elements

Approach 2: we can formulate the J-minimising solution of $Ku = f$, i.e.

$$\min_u J(u) + \alpha\| f - Ku \|_Y$$

Example: $K = I$, $J(u) = \mathrm{TV}(u)$

f

u_∞

v_∞ with $\|v_\infty\| \approx 801.88$

Computing source condition elements

Approach 2: we can formulate the J-minimising solution of $Ku = f$, i.e.

$$\min_u J(u) + \alpha \| f - Ku \|_Y$$

Example: $K = I$, $J(u) = \mathrm{TV}(u)$

f

Reconstructed image

u_α

Source condition element

v_α with $\|v_\alpha\| = \alpha = 100$

Computing source condition elements

Approach 2: we can formulate the J-minimising solution of $Ku = f$, i.e.

$$\min_u J(u) + H(f - Ku)$$

as a primal-dual problem

$$\min_u \sup_v J(u) + \langle v, f - Ku \rangle - H^\star(v)$$

Can (for example) be solved with primal-dual hybrid gradient* method, i.e.

$$u^{k+1} = \mathrm{prox}_{\tau J}\left(u^k + \tau K^* v^k\right)$$
$$v^{k+1} = \mathrm{prox}_{\sigma H^\star}\left(v^k - \sigma\left(K(2u^{k+1} - u^k) - f\right)\right)$$

$$\tau\sigma < 1/\|K\|^2$$

*(cf. Chambolle, A., & Pock, T. (2016). An introduction to continuous optimization for imaging. *Acta Numerica*, 25, 161-319.

Numerical results

(a) Shepp-Logan phantom

(b) Eileen Collins

Numerical results

Inverse problem

$$S\mathscr{F}u^\dagger = f$$

Subsampling operator

with discrete 2D Fourier transform $$(\mathscr{F}u)_{pq} = \frac{1}{\sqrt{n_x n_y}} \sum_{l=0}^{n_y-1} \sum_{j=0}^{n_x-1} u_{lj}\, e^{-i\frac{2\pi pl}{n_y}} e^{-i\frac{2\pi qj}{n_x}}$$

Regularisation function:

$$J(u) = H(Au) \quad \text{with} \quad (Au)_{ijp} = \begin{cases} u_{(i+1)j} - u_{ij} & p = 1 \\ u_{i(j+1)} - u_{ij} & p = 2 \end{cases}, \quad H(q) = \sum_{i=1}^{n_y-1} \sum_{j=1}^{n_x-1} \sqrt{\left|q_{ij1}\right|^2 + \left|q_{ij2}\right|^2}\,.$$

(2D discretised isotropic total variation)

Numerical results

Inverse problem

$$S\mathscr{F}u^\dagger = f$$

Subsampling operator

Source condition

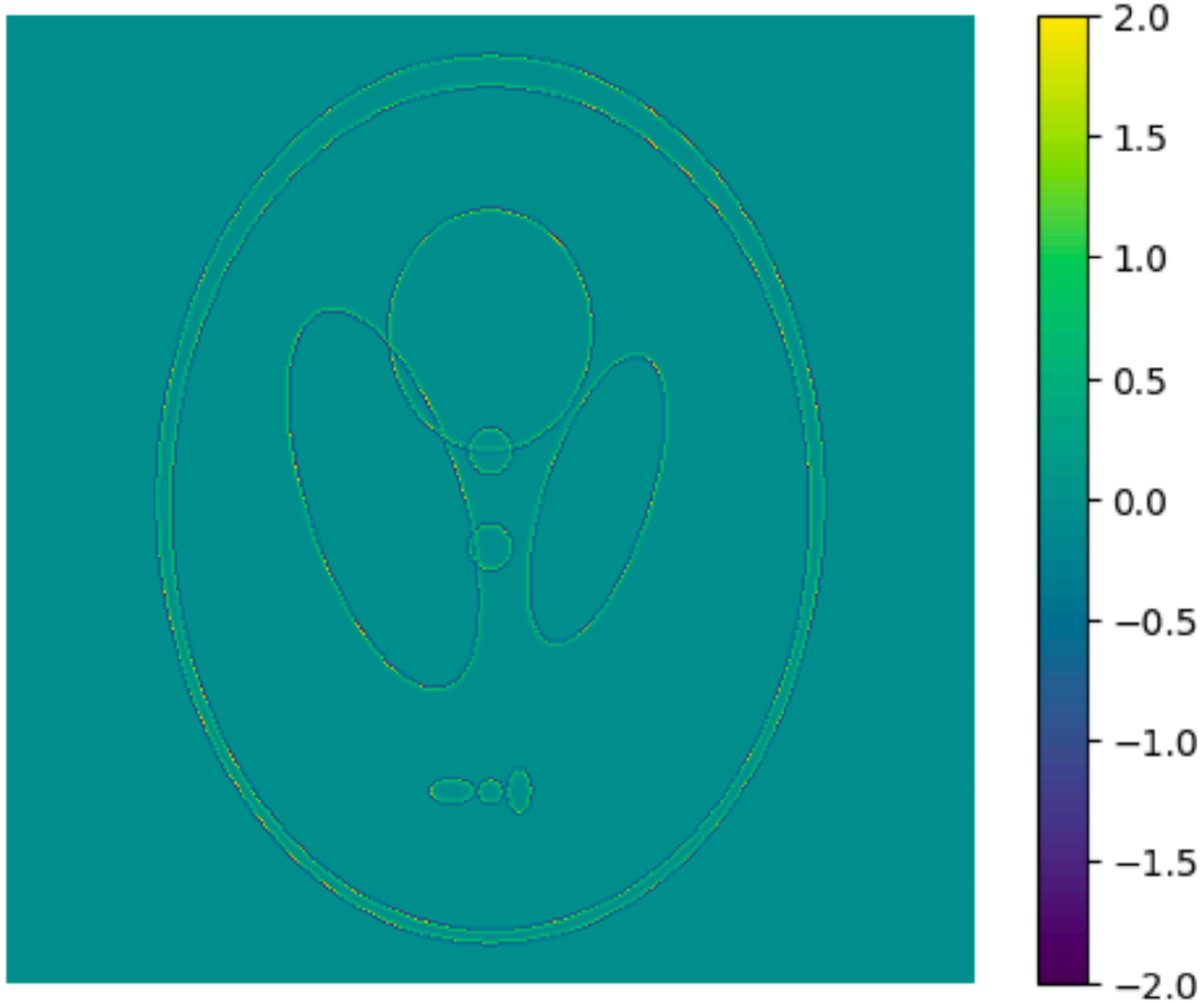
$$\mathscr{F}^{-1}S^\top v \in \partial \mathrm{TV}(u^\dagger)$$

Algorithm

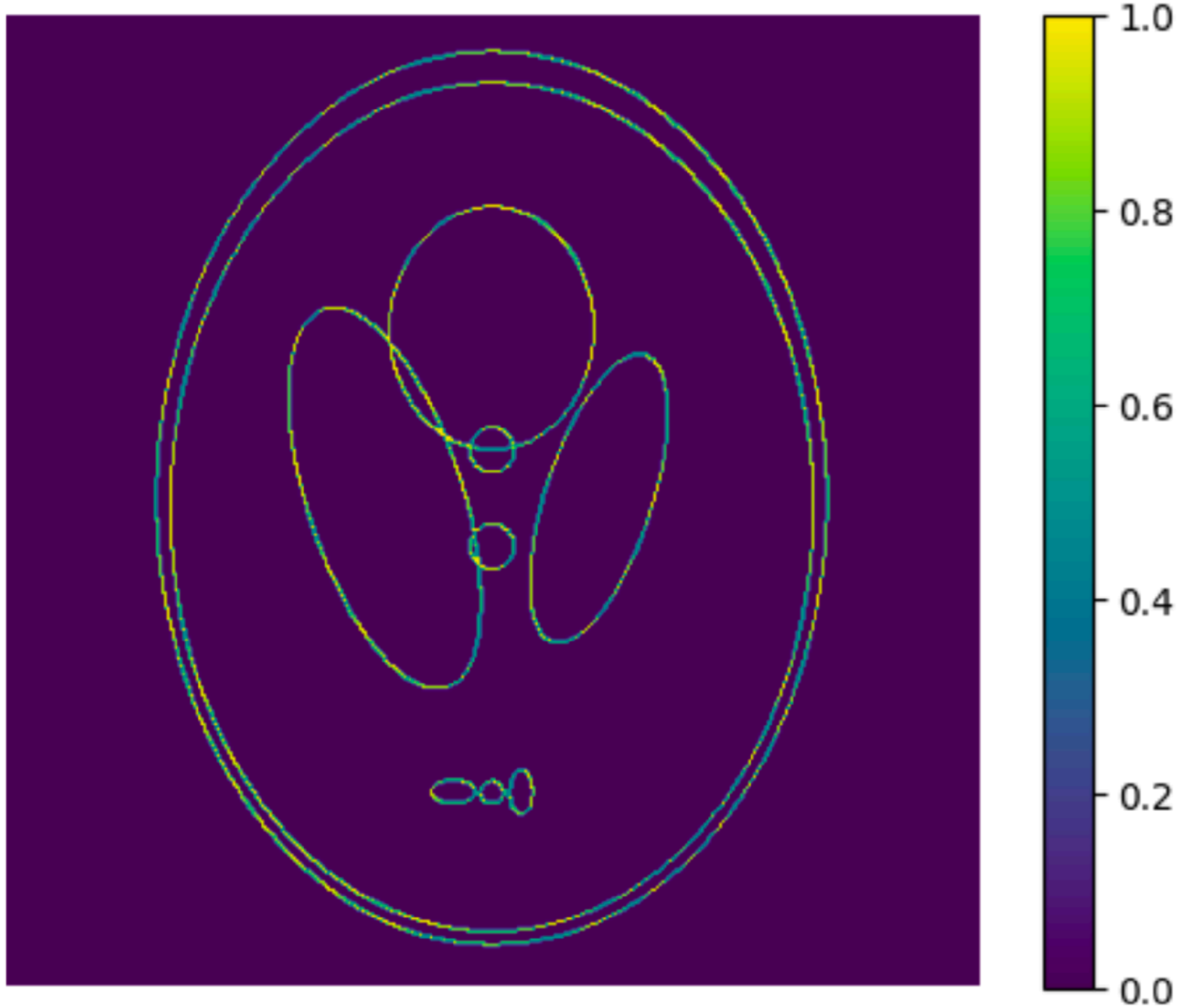
$$v^{k+1} = v^k - \tau S\mathscr{F}\left(\mathscr{F}^{-1}S^\top v^k - A^\top q^k\right),$$

$$q^{k+1} = q^k - \sigma\left(A\left(A^\top q^k - \mathscr{F}^{-1}S^\top v^{k+1}\right) + \mathrm{prox}_{\|\cdot\|_{2,1}}\left(Au^\dagger + q^k\right) - Au^\dagger\right).$$

Numerical results



$S = I$

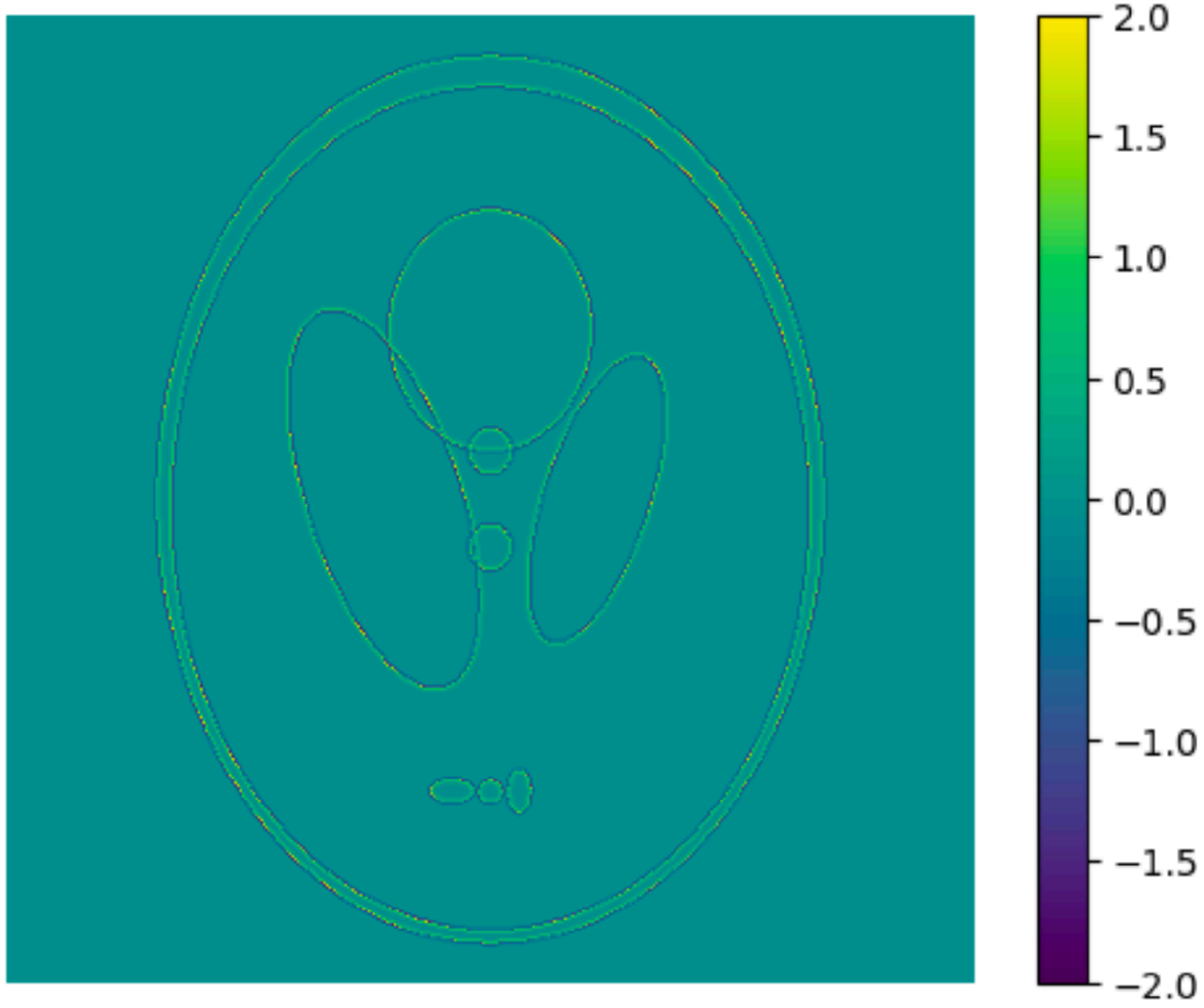


(a) $\mathscr{F}^{-1}v^K$

$\mathscr{F}^{-1}v^K = A^\top q^K$

(b) $\sqrt{|q_1^K|^2 + |q_2^K|^2}$

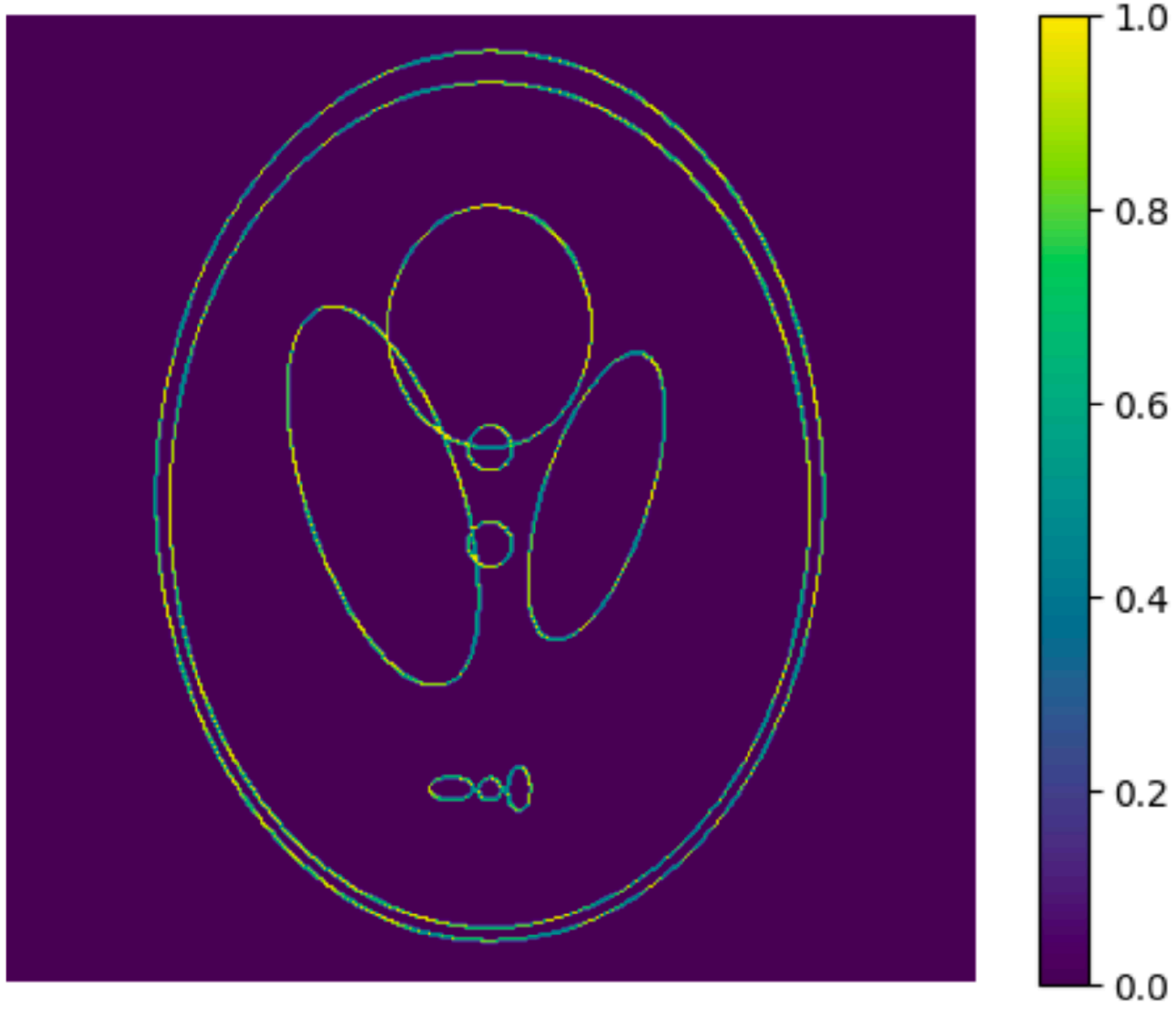
Numerical results



(a) $\mathscr{F}^{-1}v^K$

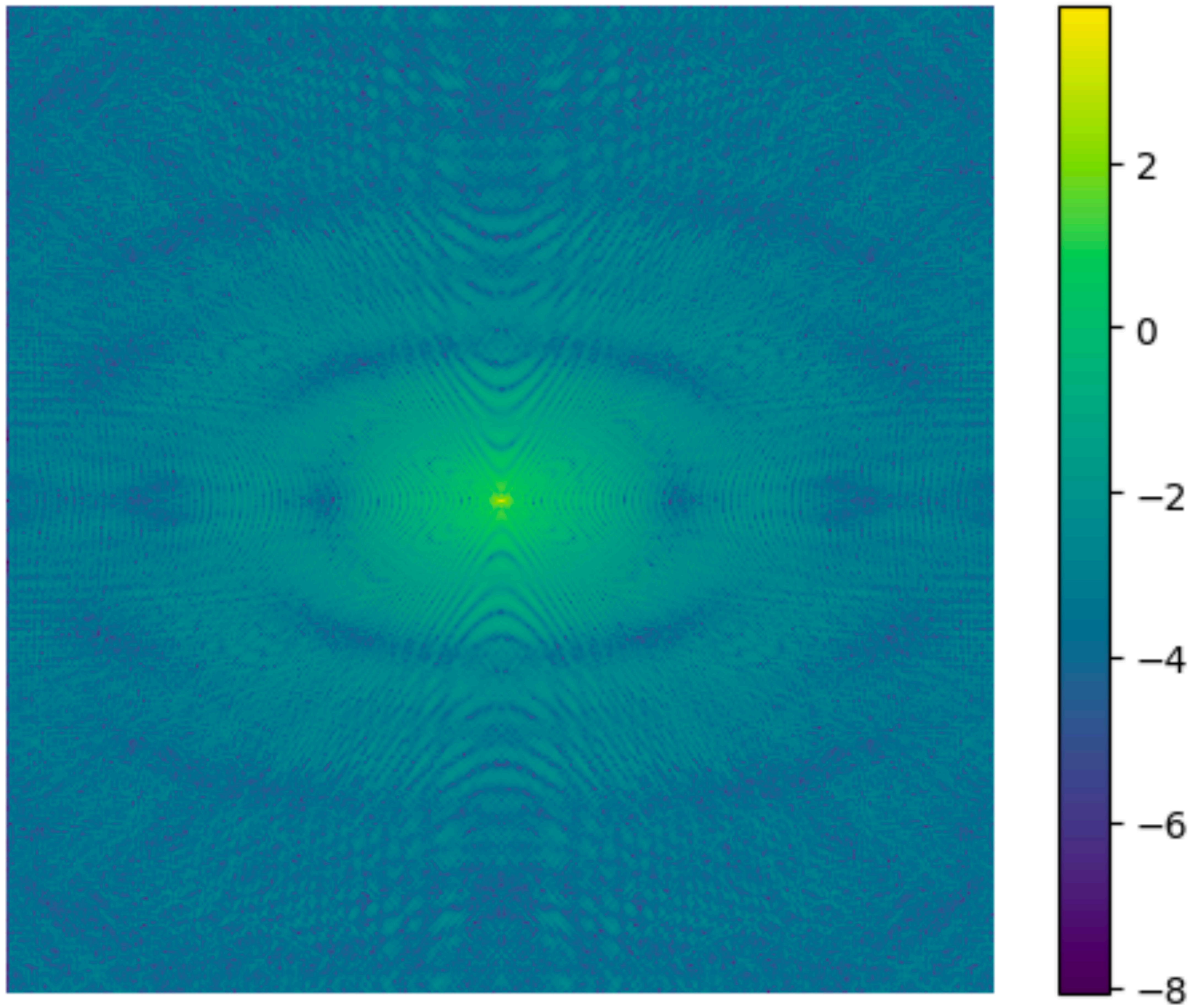
$S = I$

$\|v^K\| \approx 101.78$

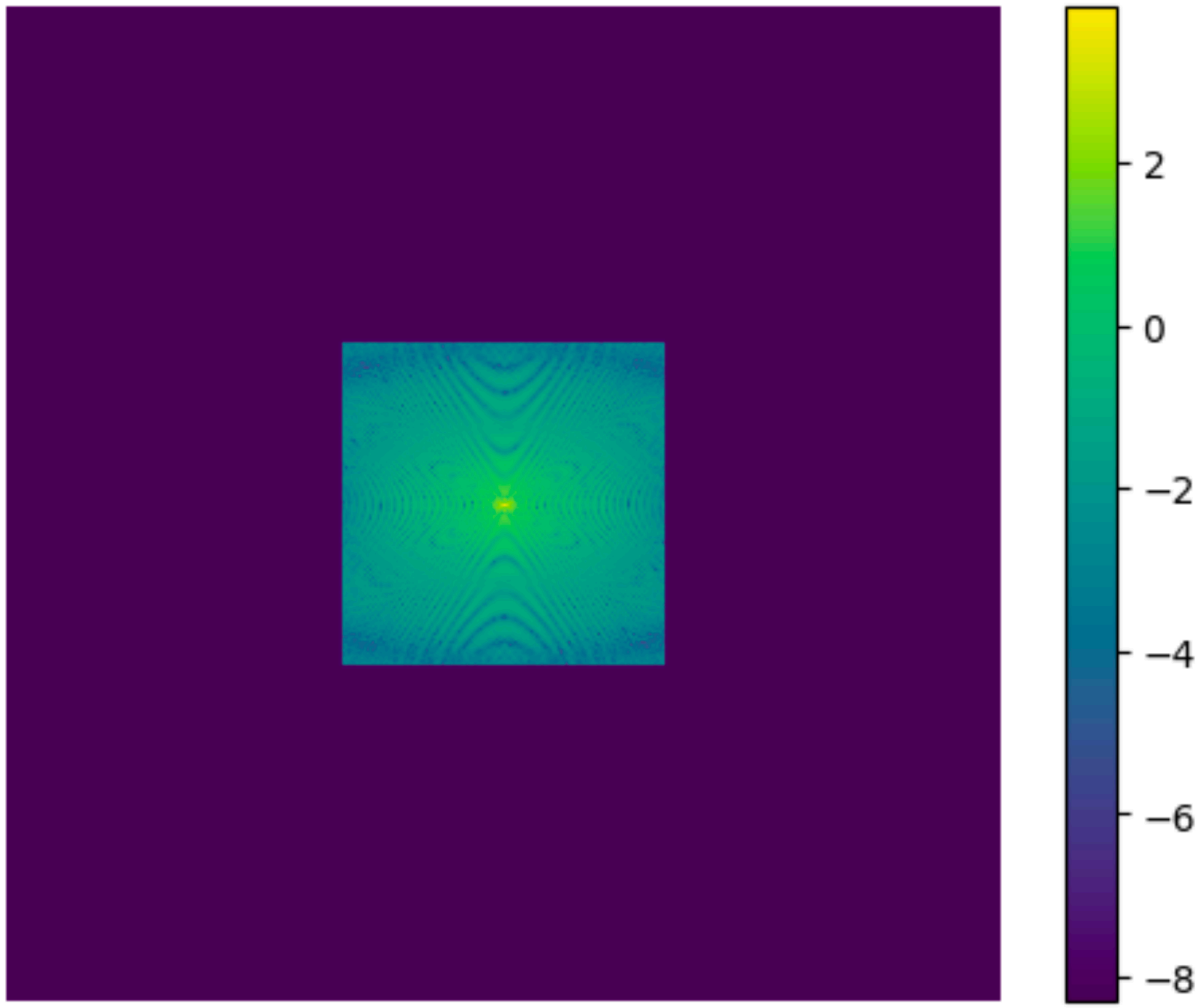


(b) $\sqrt{|q_1^K|^2 + |q_2^K|^2}$

Numerical results

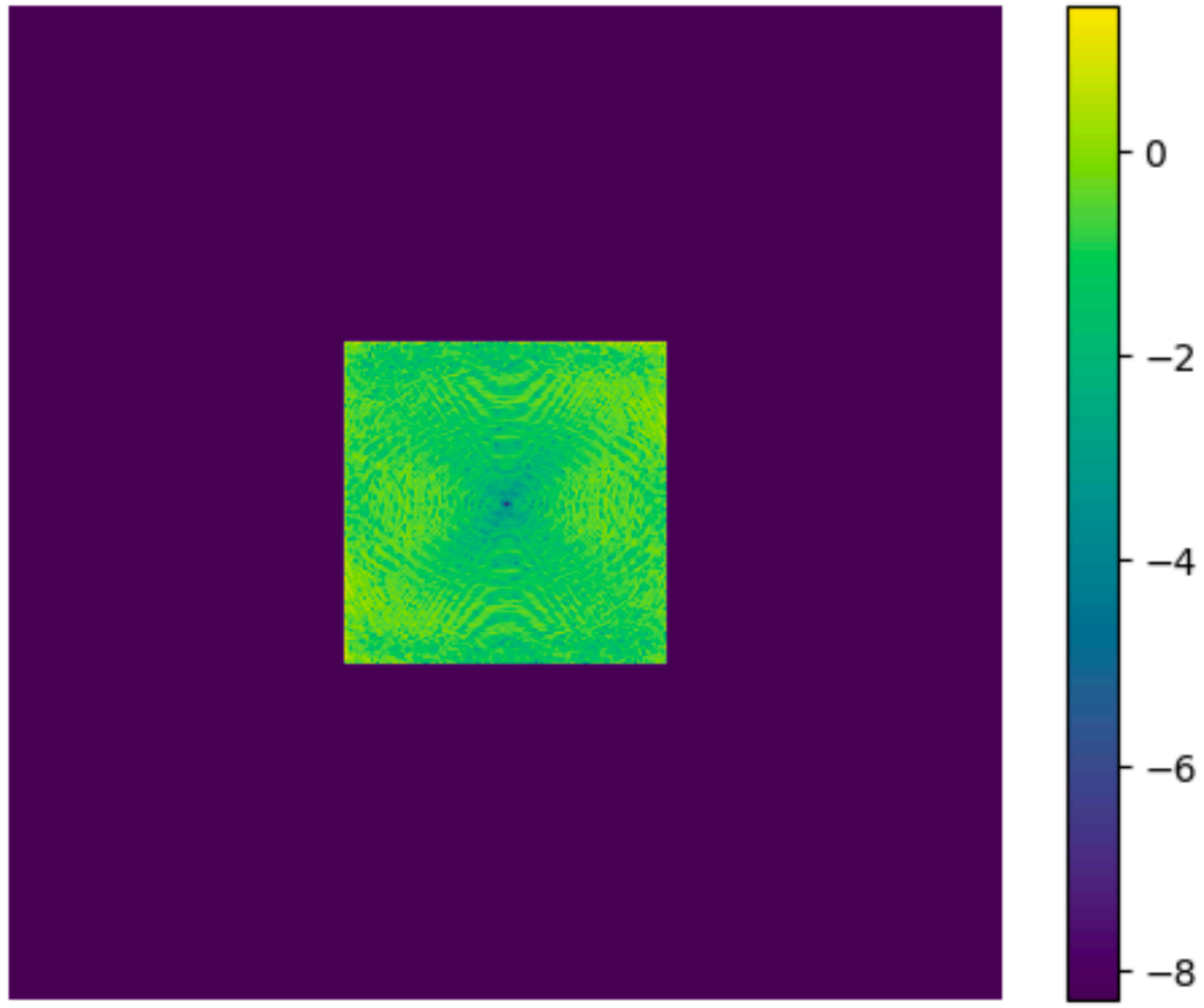


(a) Fourier transformed data

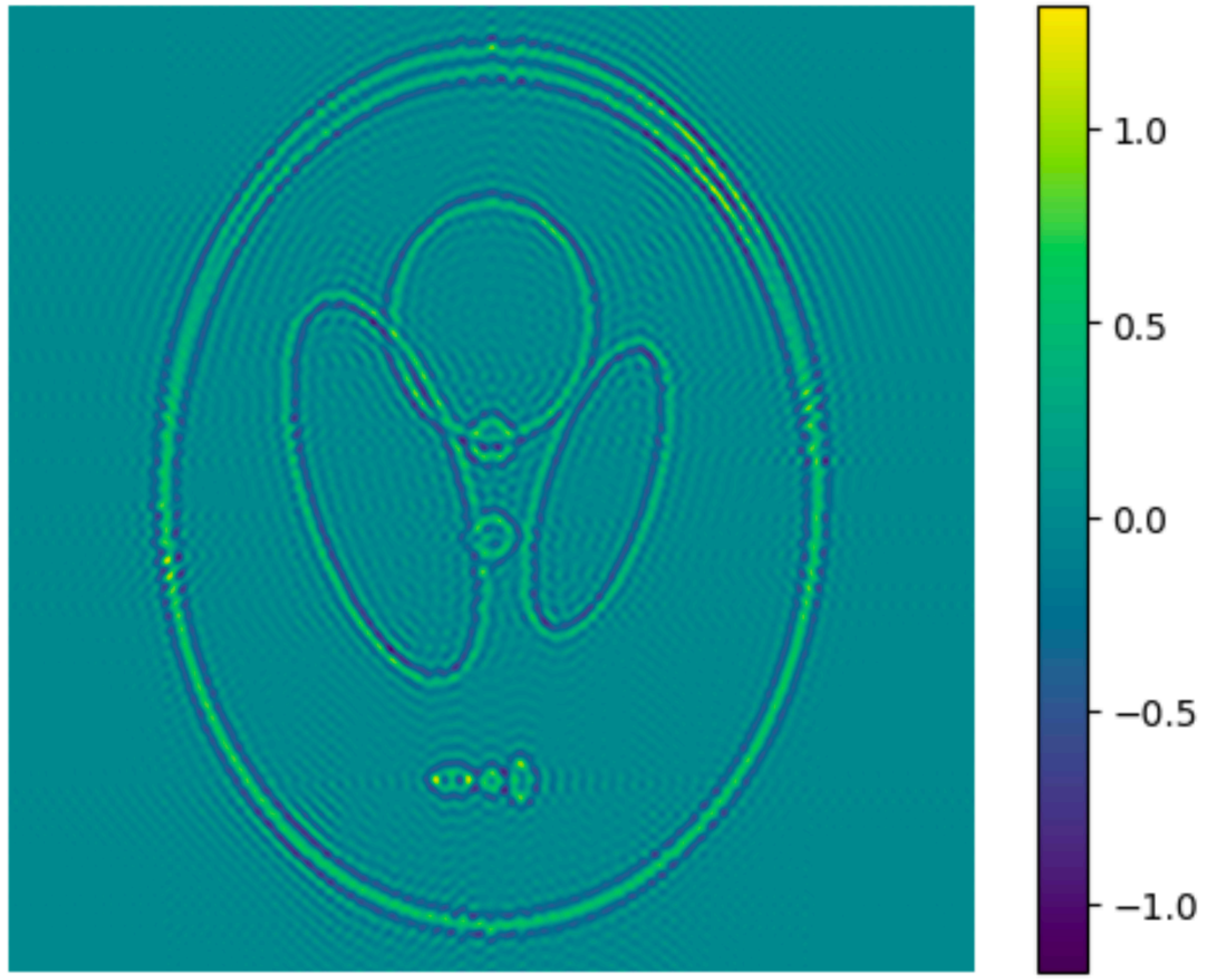


(b) Sub-sampled Fourier transformed data

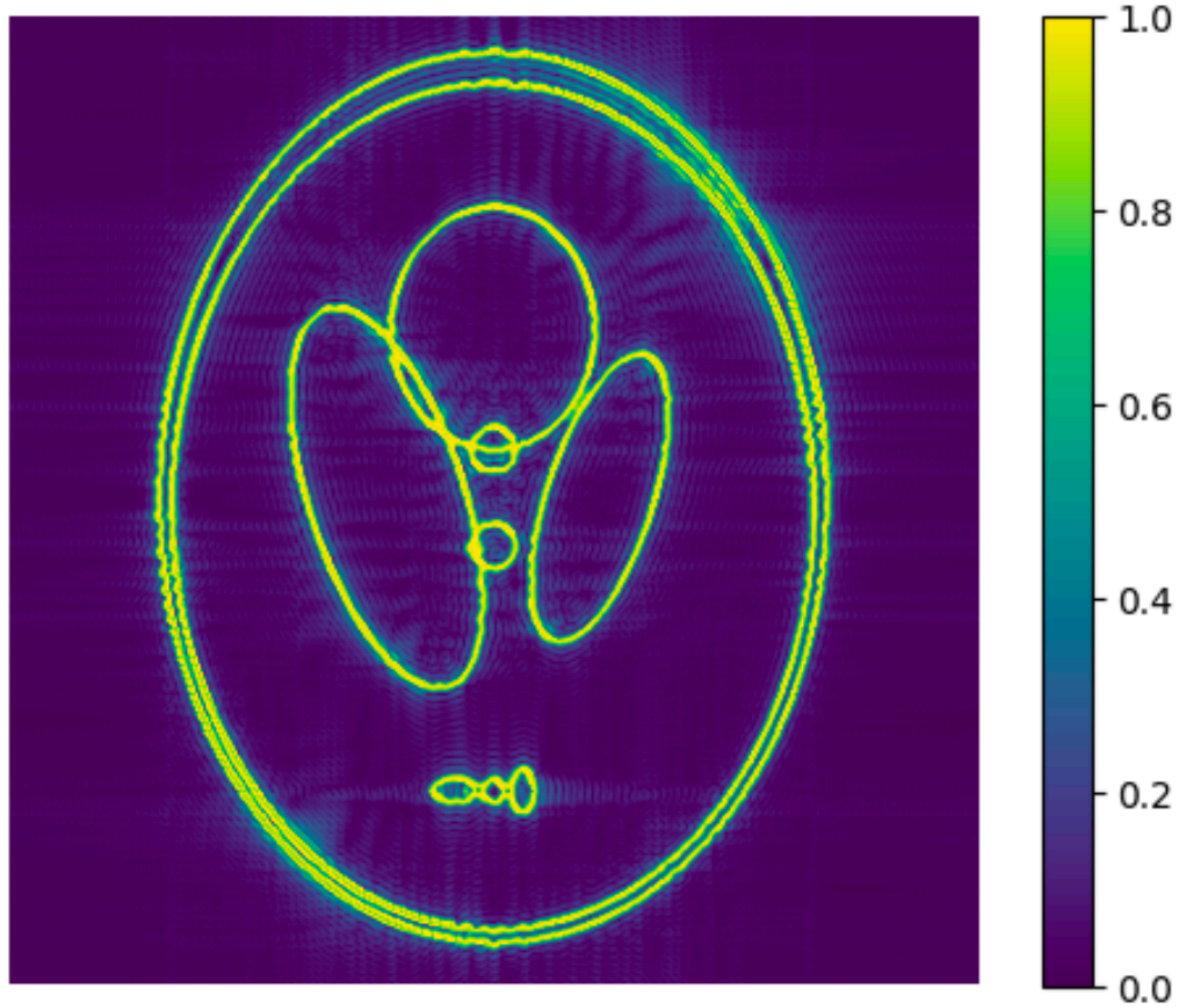
Numerical results



(a) v^K



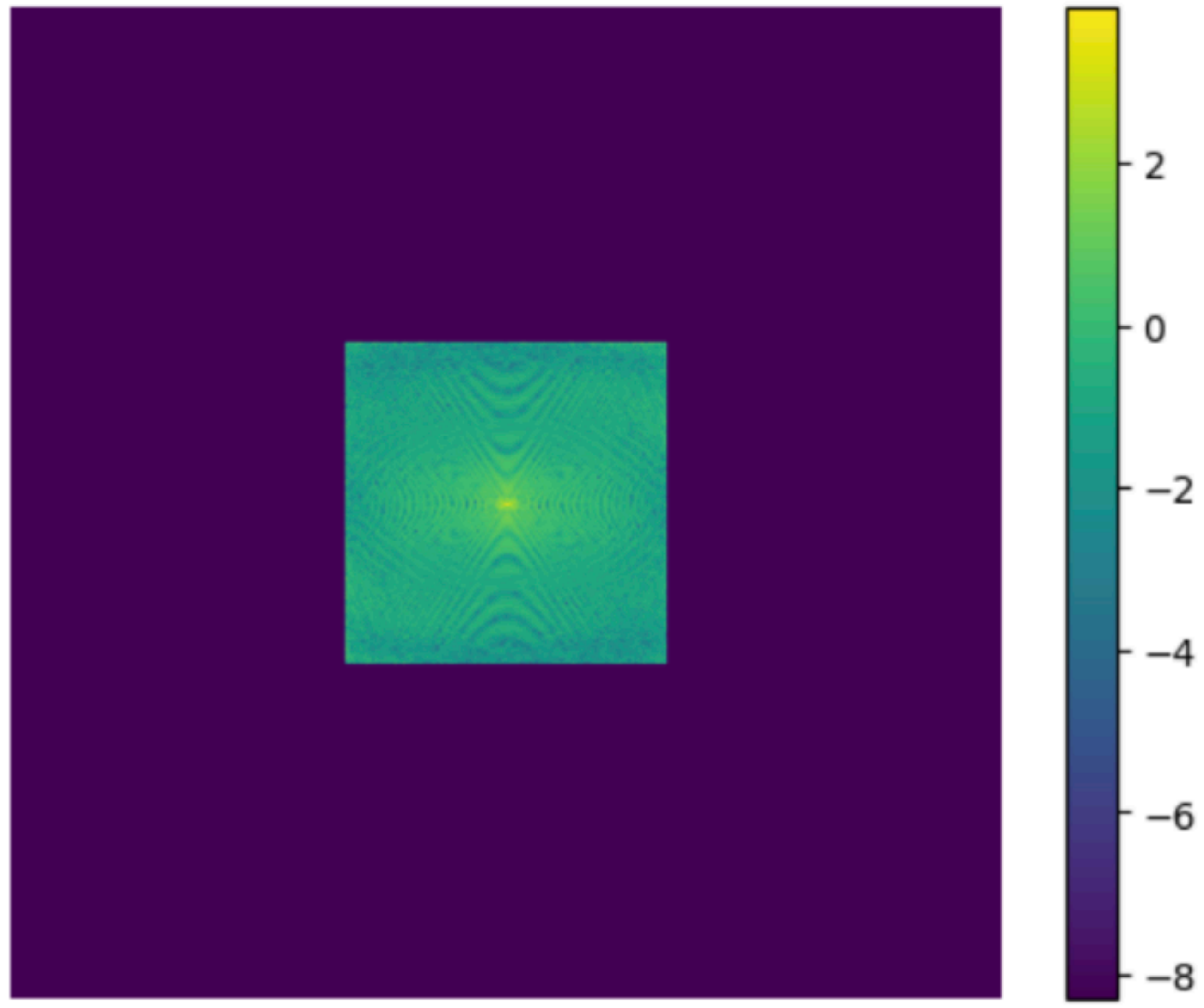
(b) Backprojection $\mathcal{F}^{-1} S^\top v^K$



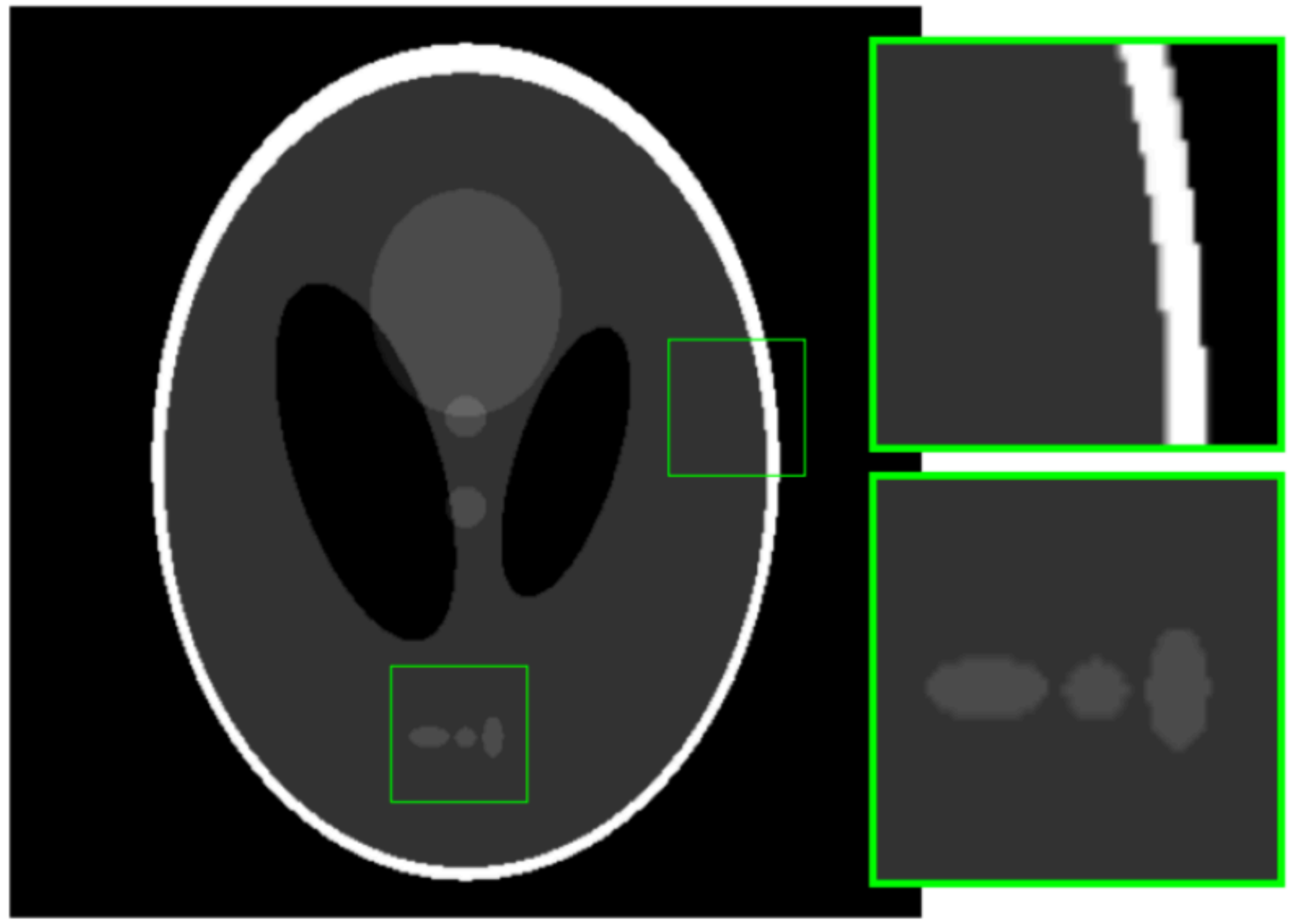
(c) $\sqrt{|q_1^K|^2 + |q_2^K|^2}$

$$\mathscr{F}^{-1} S^T v^K = A^\top q^K$$

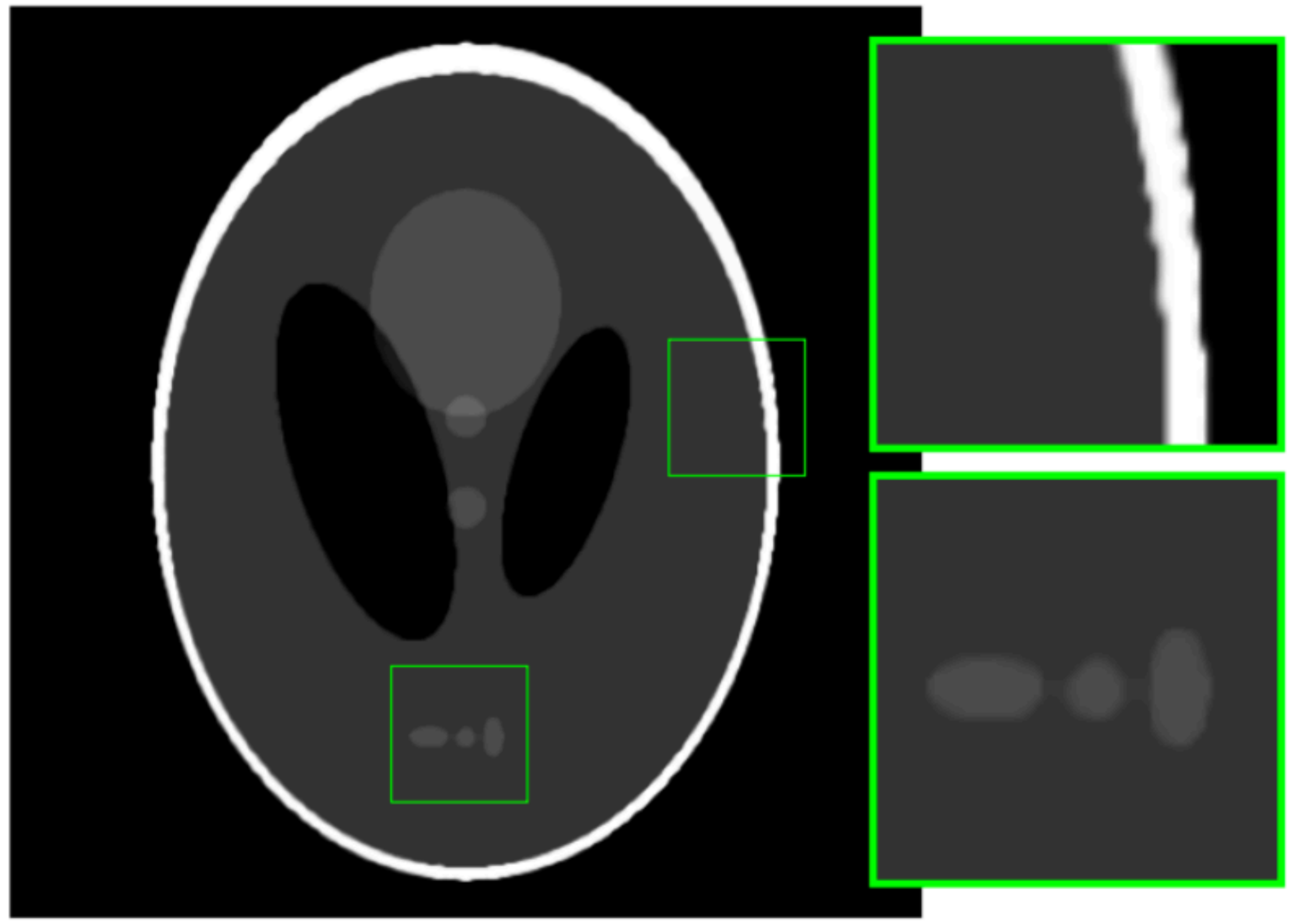
$$\|v^K\| \approx 72.79$$



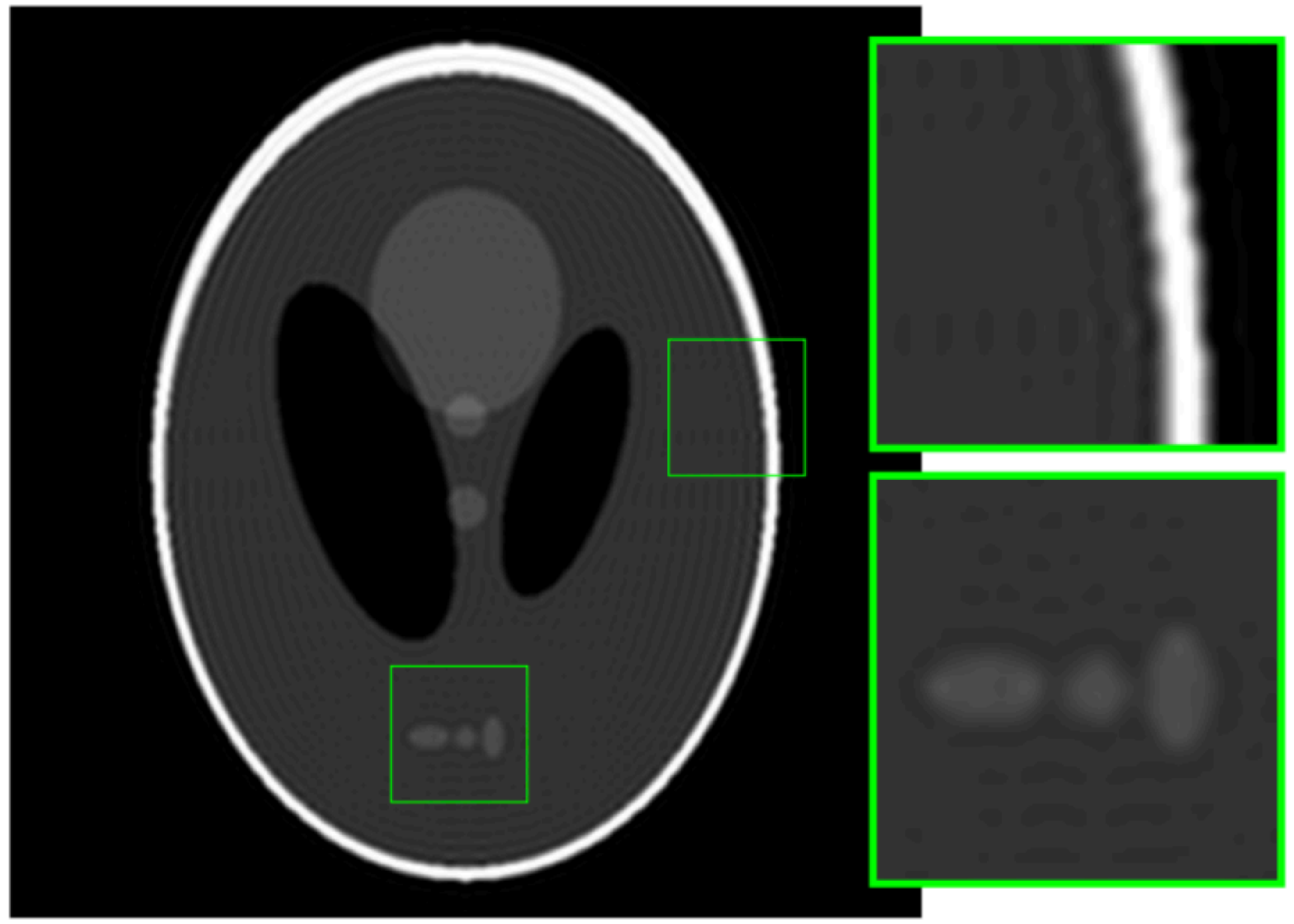
(a) Range cond. data g_α^K



(b) ground truth $u^\dagger$



(c) approximate solution u^N *



(d) low-pass filtered rec. $\mathcal{F}^{-1}S^\top S\mathcal{F}u^\dagger$

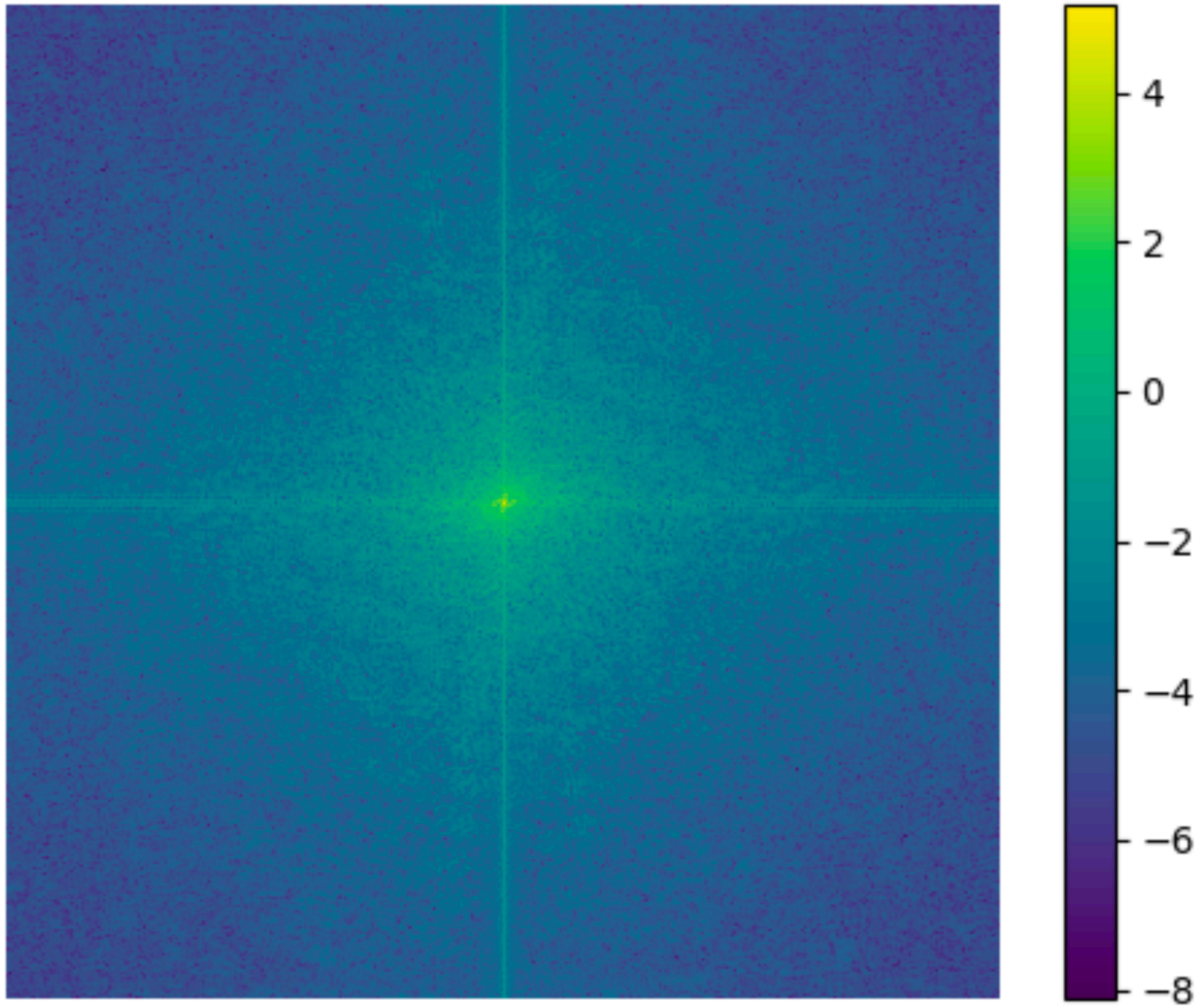
*Sanity check computed with PDHG method (cf. Chambolle, A., & Pock, T. (2016). An introduction to continuous optimization for imaging. *Acta Numerica*, *25*, 161-319.)

Numerical results

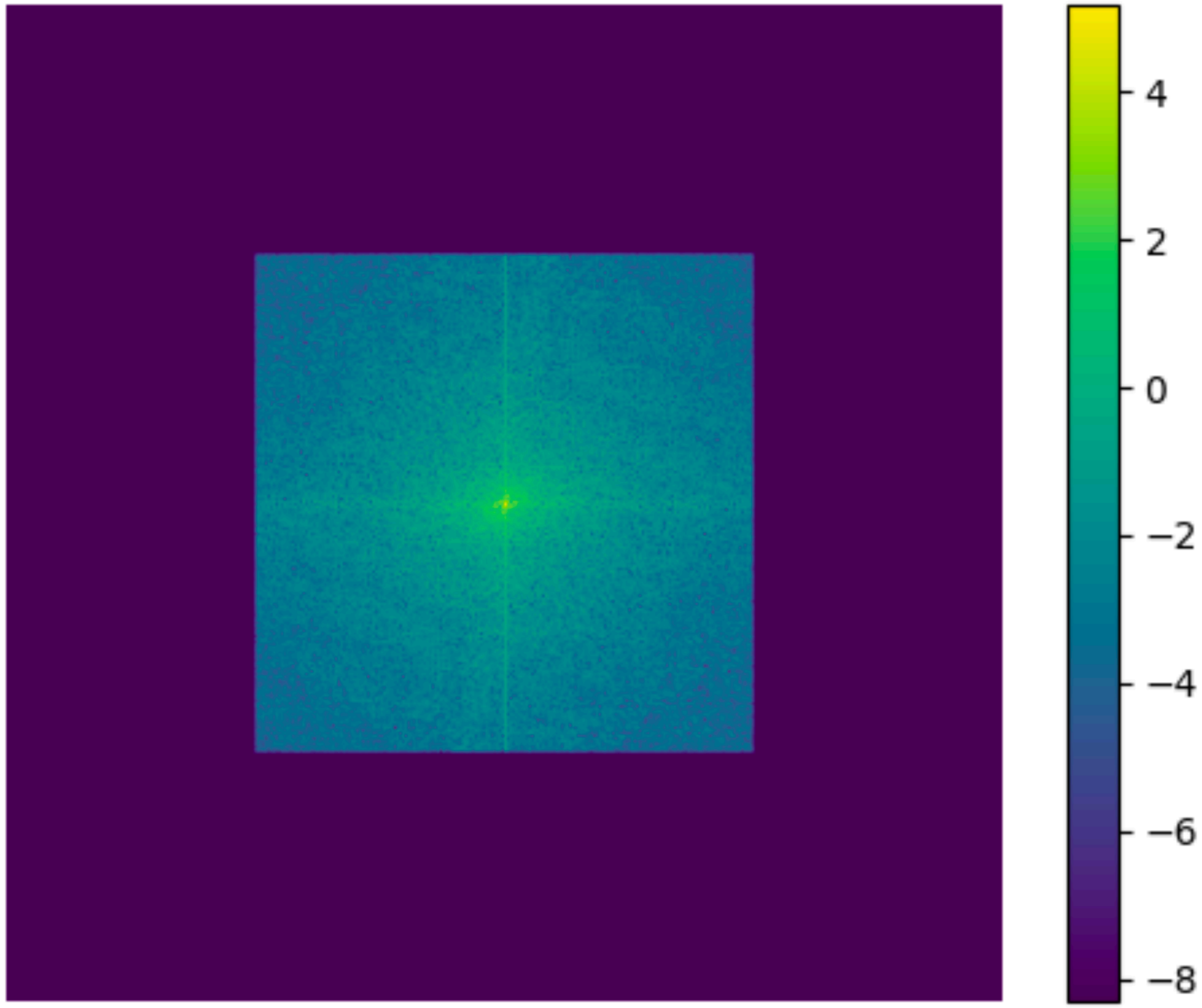
(a) Shepp-Logan phantom

(b) Eileen Collins

Numerical results

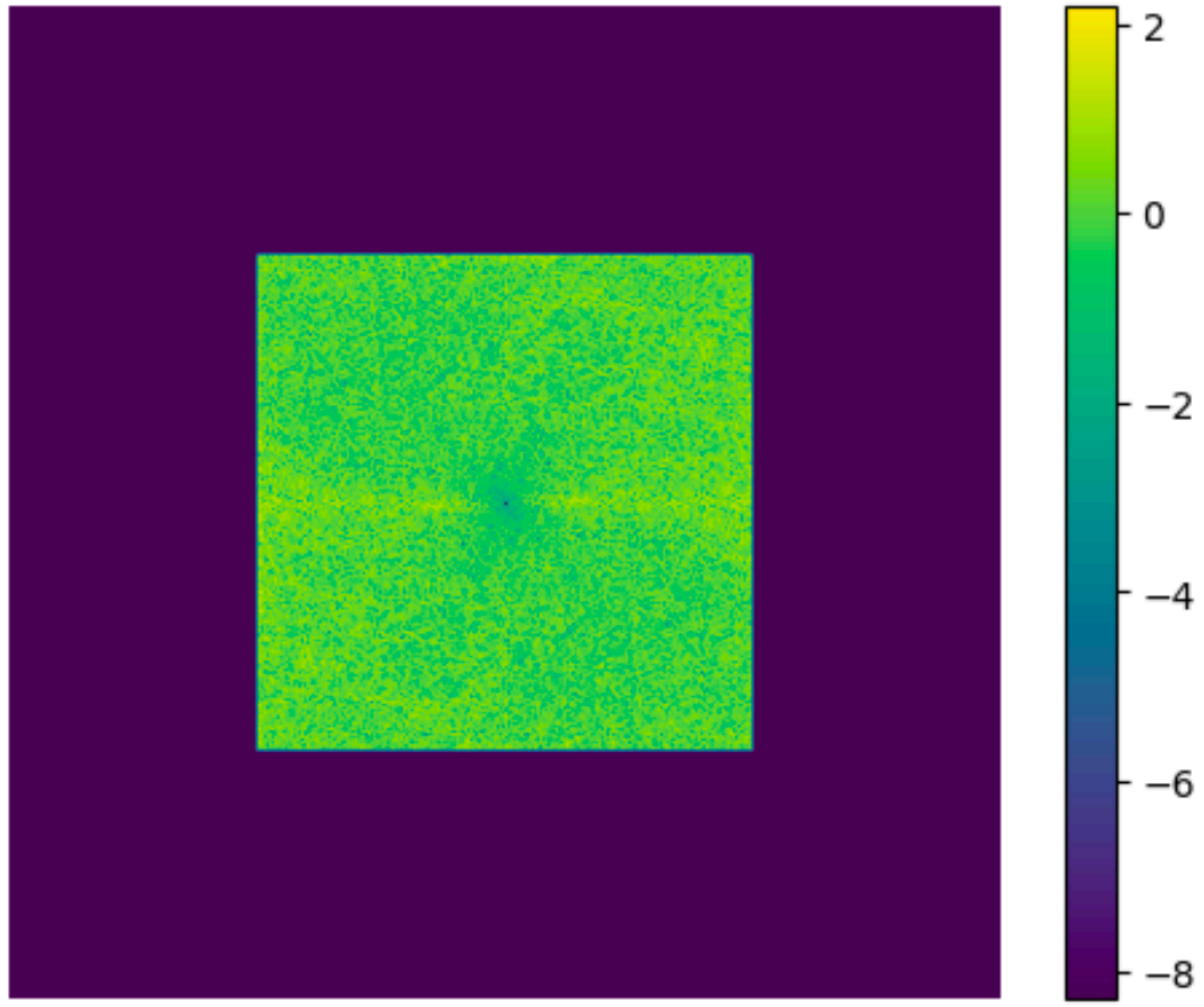


(a) Fourier transformed data

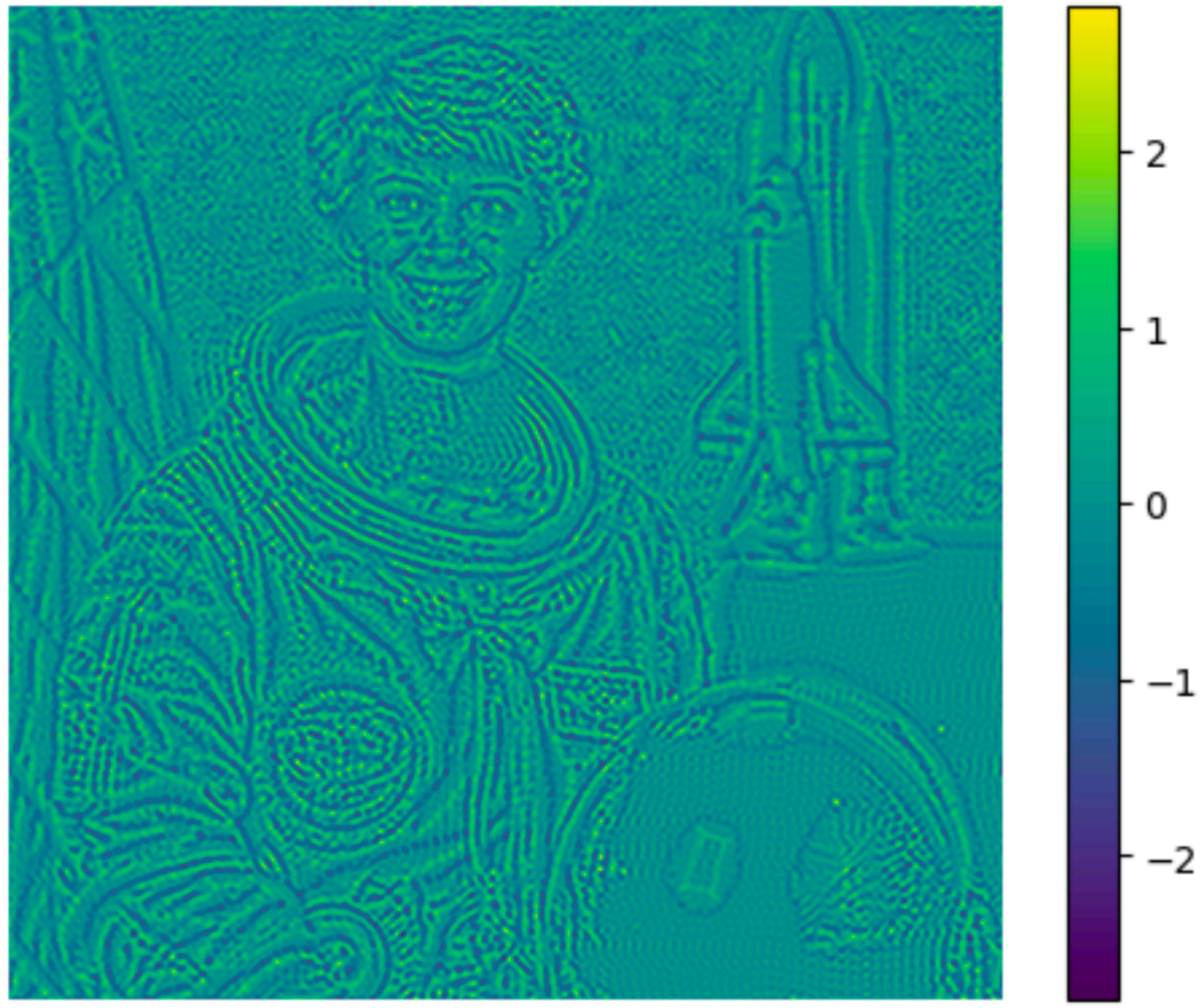


(b) Sub-sampled Fourier transformed data

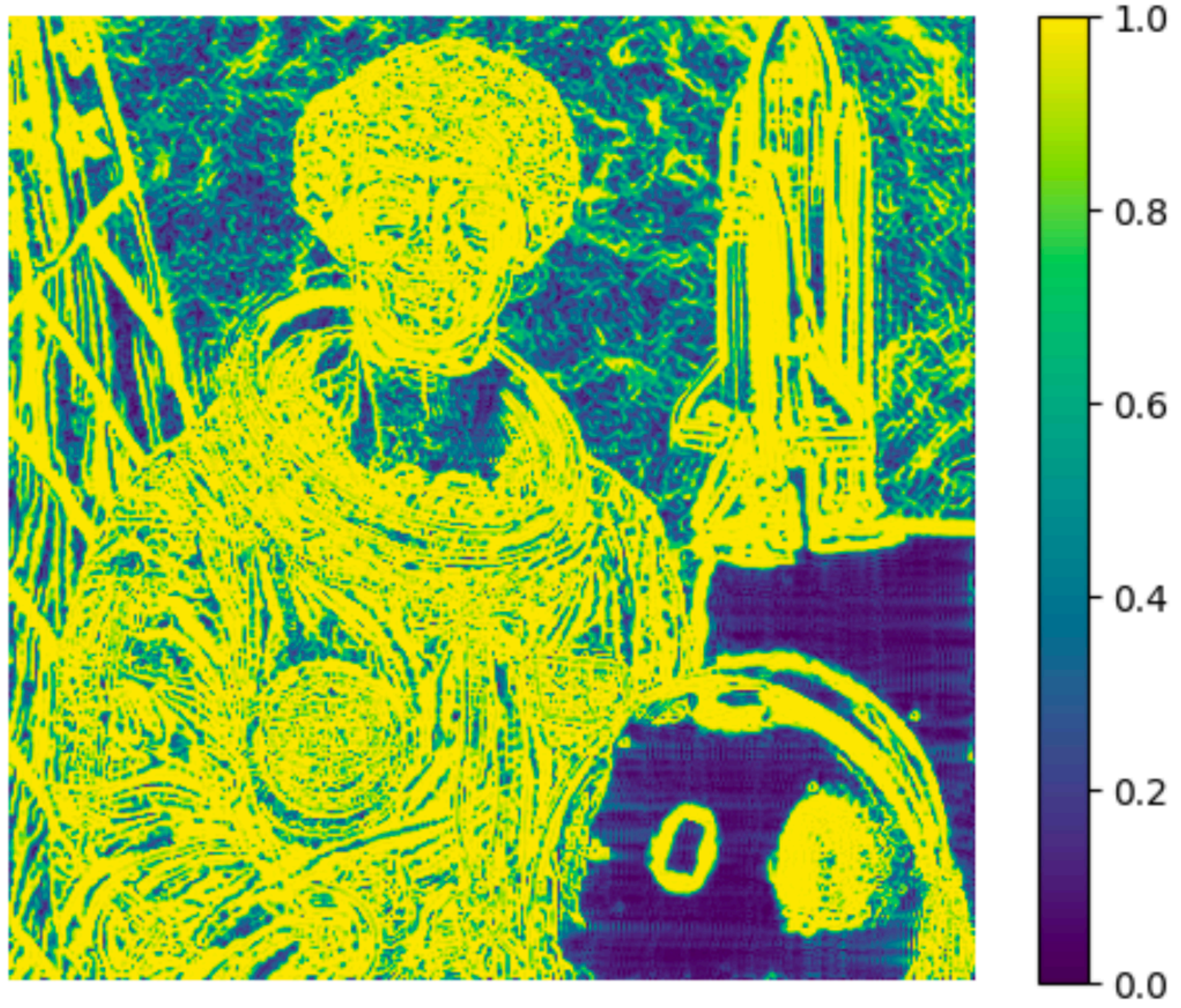
Numerical results



(a) v^K



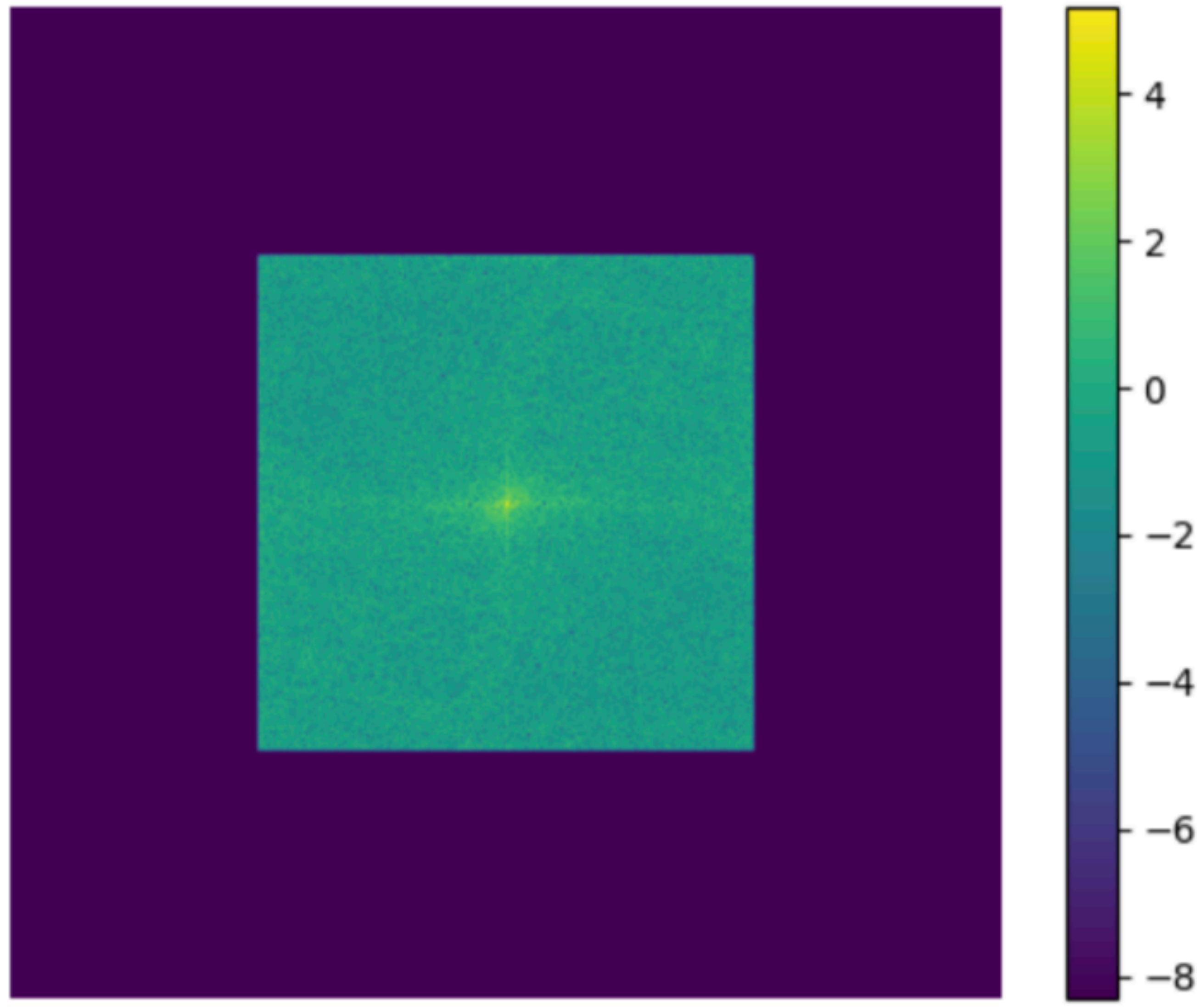
(b) Backprojection $\mathcal{F}^{-1} S^\top v^K$



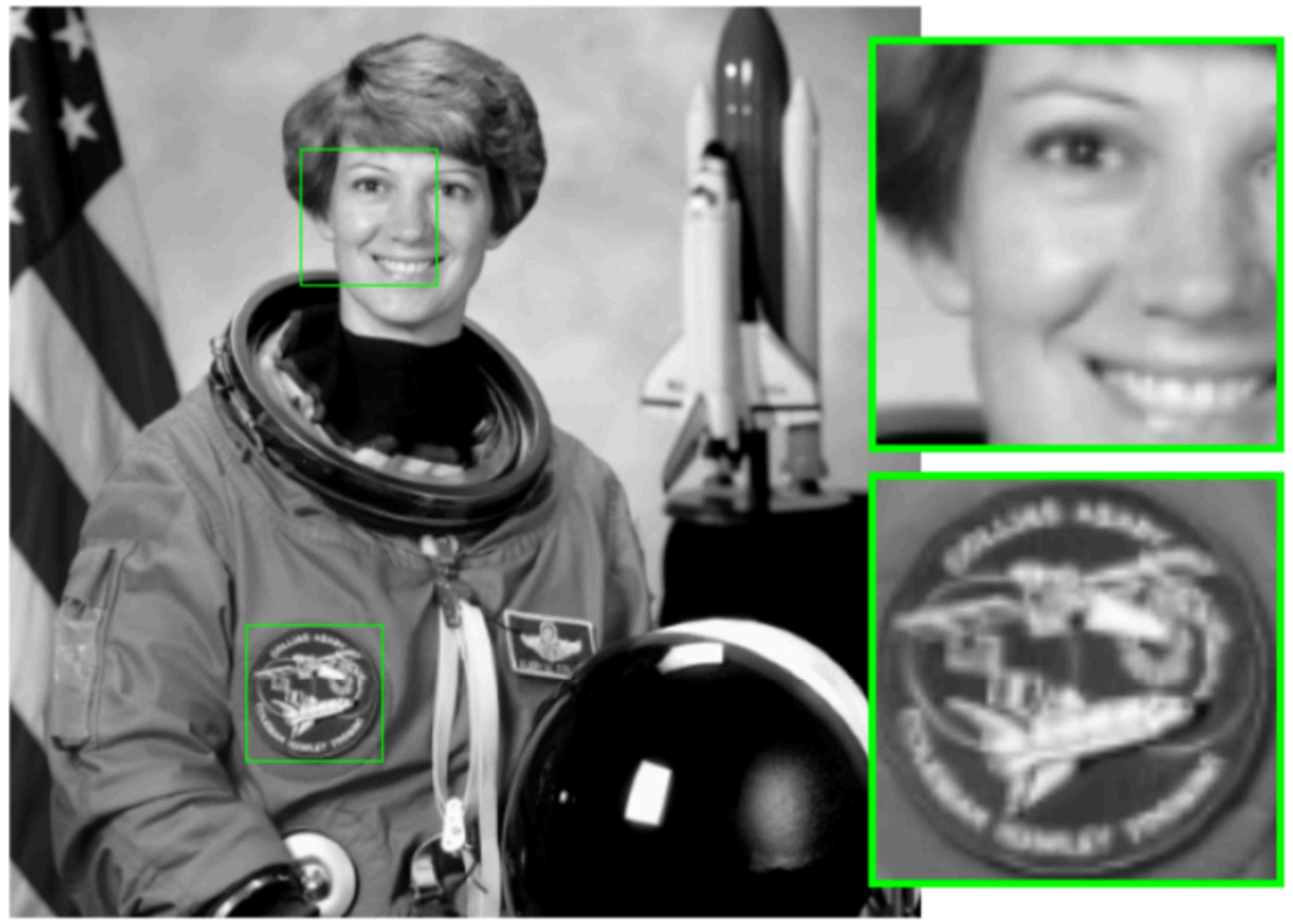
(c) $\sqrt{|q_1^K|^2 + |q_2^K|^2}$

$$\mathscr{F}^{-1} S^T v^K = A^\top q^K$$

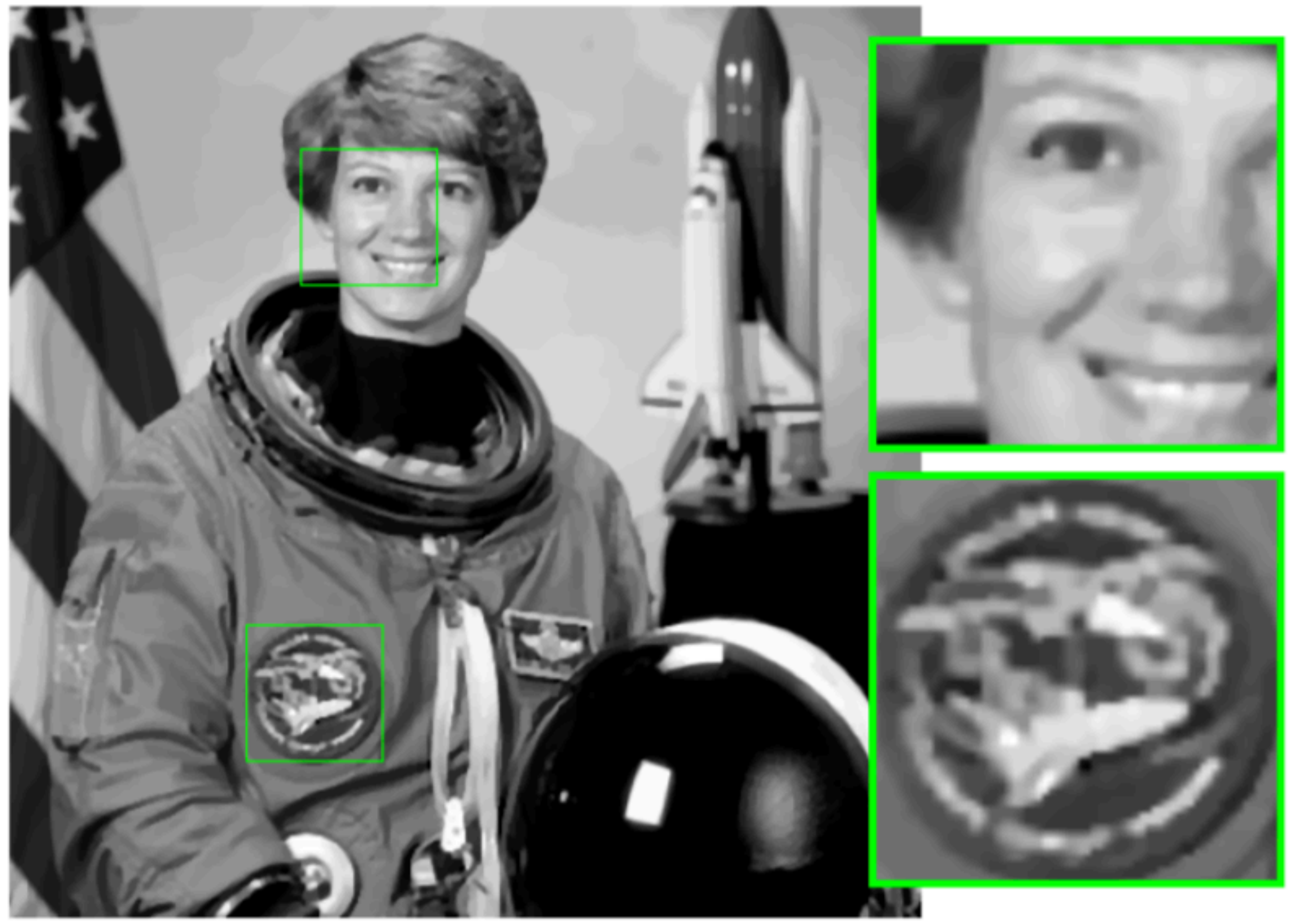
$$\|v^K\| \approx 255.15$$



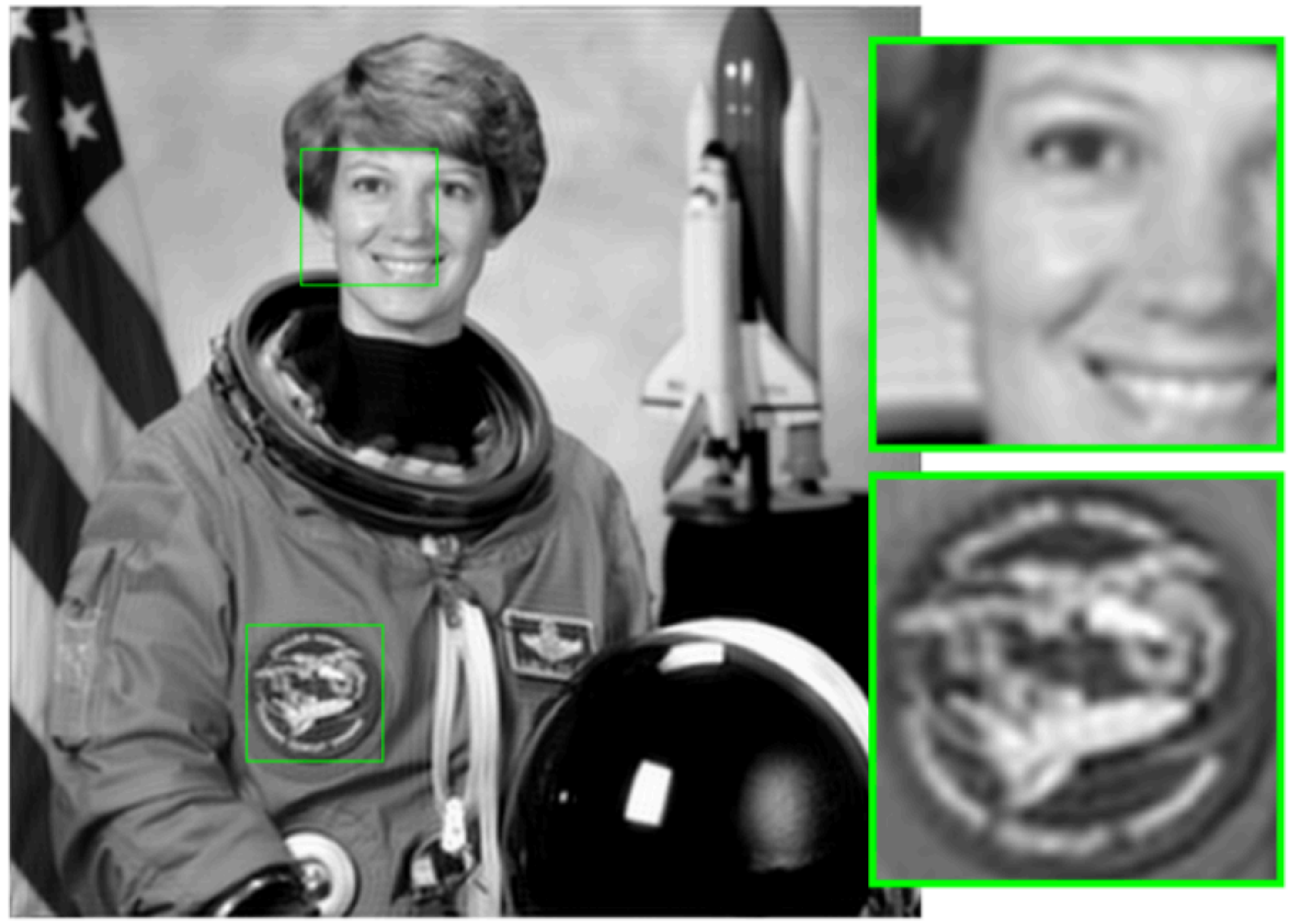
(a) Range cond. data g_α^K



(b) ground truth $u^\dagger$



(c) approximate solution u^N *



(d) low-pass filtered rec. $\mathcal{F}^{-1}S^\top S\mathcal{F}u^\dagger$

*Sanity check computed with PDHG method (cf. Chambolle, A., & Pock, T. (2016). An introduction to continuous optimization for imaging. *Acta Numerica*, *25*, 161-319.)

Learning variational regularisations with optimal error estimates

We could approximate a bilevel problem with trainable regularisation function

$$\min_{\Theta} \frac{1}{2s} \sum_{i=1}^{s} \|u(\Theta) - u^\dagger\|_X^2 + \beta R(\Theta)$$

subject to

$$u(\Theta) \in \arg\min_u \left\{ \frac{1}{2}\|Ku - f^\delta\|_Y^2 + J(u,\Theta) \right\}$$ (Lower level problem)

and check if learned regularisation function $J(\,\cdot\,,\Theta)$ improves norm of source condition elements

Choose $R(\Theta) = 0$ $\quad J(u,A) = \sum_{j=1}^{n} \sqrt{\sum_{l=1}^{m} (Au)_{jl}}$ $\quad K = I$

Approximate lower level problem with finite number of PDHG iterations

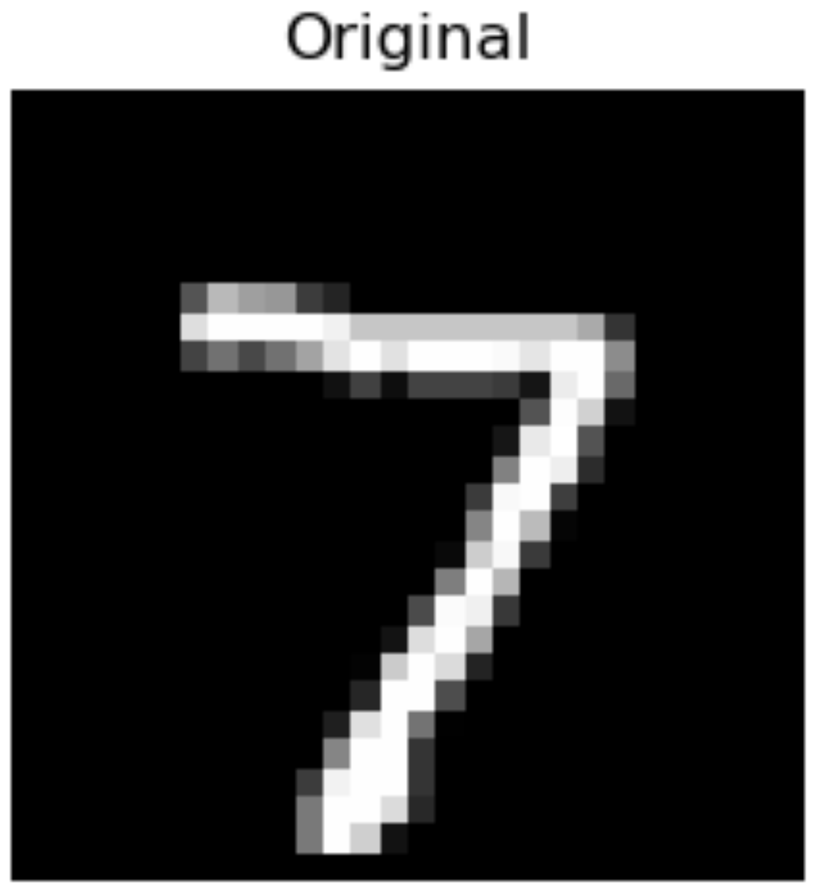
Choose $R(\Theta) = 0$ $J(u, A) = \sum_{j=1}^{n} \sqrt{\sum_{l=1}^{m} (Au)_{jl}}$ $K = I$

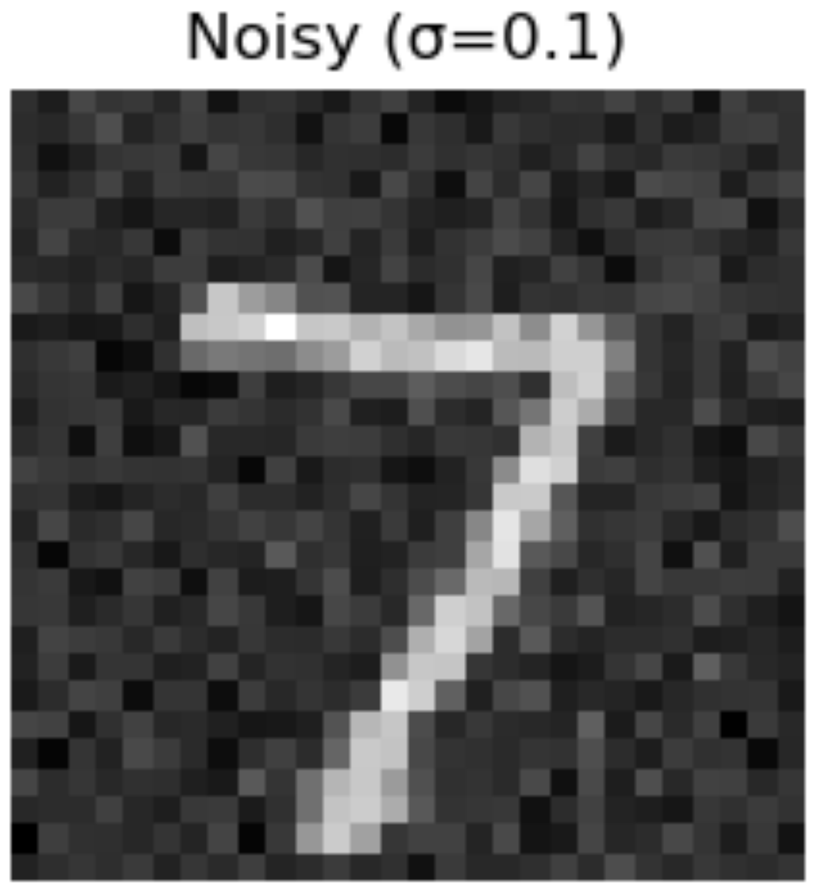
Approximate lower level problem with finite number of PDHG iterations

Example: convolution $Au = \begin{pmatrix} h_1 \\ \vdots \\ h_m \end{pmatrix} * u$

Samples

$u_i^\dagger =$





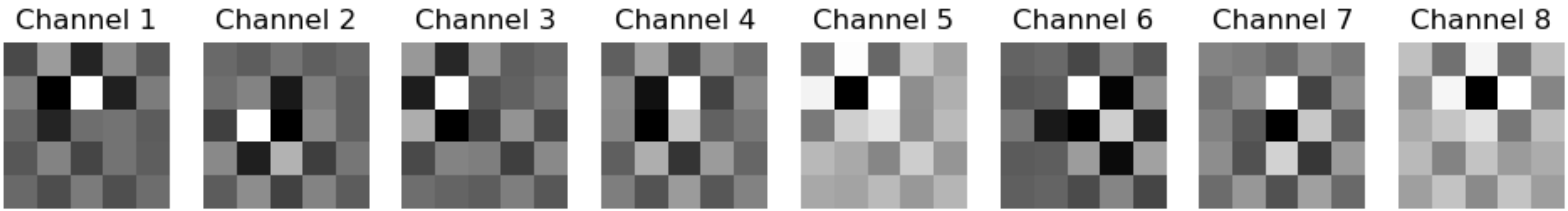
$= f_i^\delta$

Choose $R(\Theta) = 0$ $J(u, A) = \sum_{j=1}^{n} \sqrt{\sum_{l=1}^{m} (Au)_{jl}}$ $K = I$

Approximate lower level problem with finite number of PDHG iterations

Example: convolution $Au = \begin{pmatrix} h_1 \\ \vdots \\ h_m \end{pmatrix} * u$

Optimal kernel ($m = 8$):



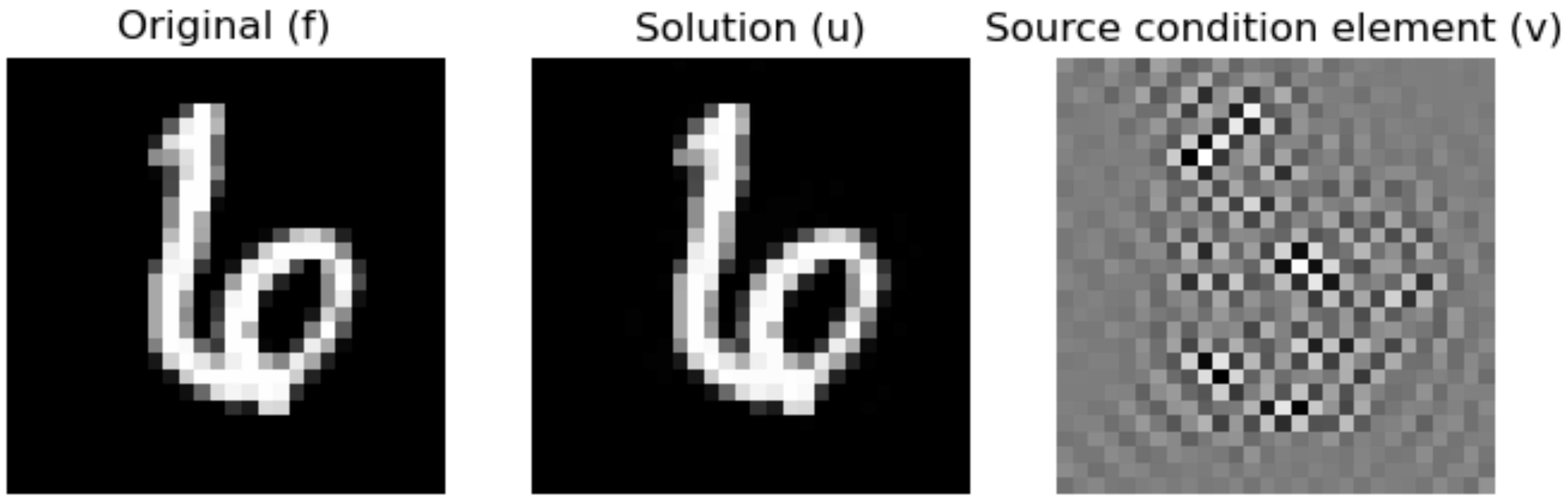
Choose $R(\Theta) = 0$ $\quad J(u, A) = \sum_{j=1}^{n} \sqrt{\sum_{l=1}^{m} (Au)_{jl}}$ $\quad K = I$

Approximate lower level problem with finite number of PDHG iterations

Example: convolution $\quad Au = \begin{pmatrix} h_1 \\ \vdots \\ h_m \end{pmatrix} * u$

$\|v\| \approx 8.59$

Source condition computation



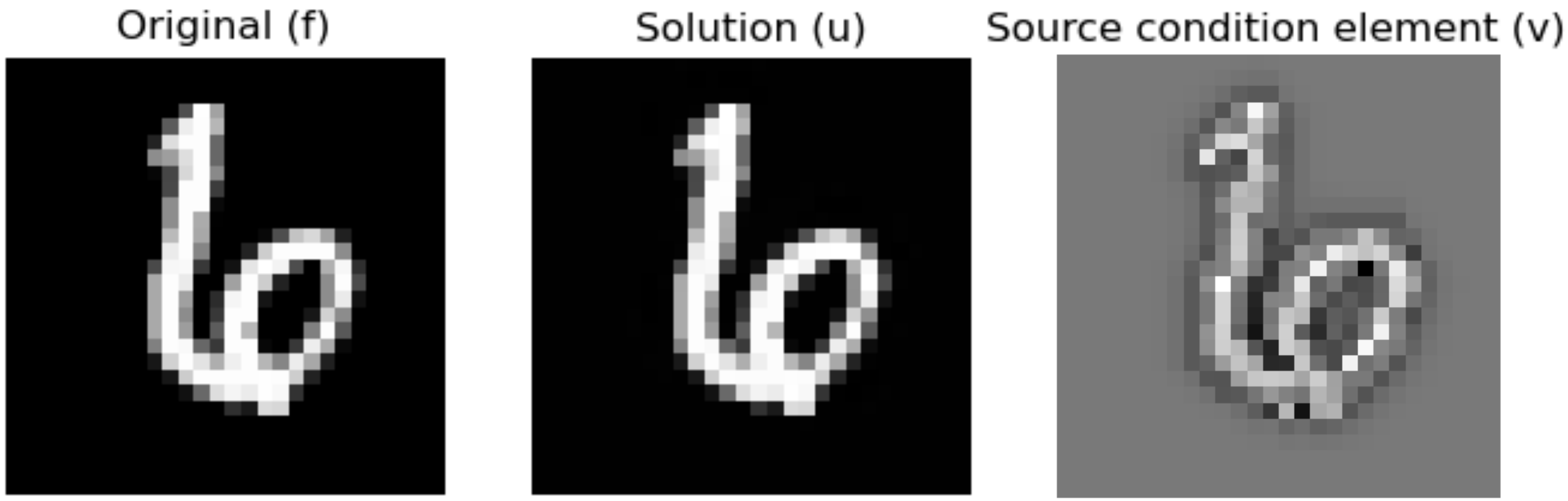
Choose $R(\Theta) = 0$ $\mathrm{TV}(u) = \sum_{j=1}^{n} \sqrt{\sum_{l=1}^{m} (\nabla u)_{jl}}$ $K = I$

Approximate lower level problem with finite number of PDHG iterations

Comparison with total variation

$\|v\| \approx 17.376$

Source condition computation



Learning the sampling pattern in MRI

Numerical results

Back to the source condition $\mathscr{F}^{-1}S^\top v \in \partial \mathrm{TV}(u^\dagger)$

Idea: define $\tilde{v} := S^\top v \in \mathbb{C}^{n_y \times n_x}$

and estimate sparse $\tilde{v}$ instead of v by solving

$$\min_{\tilde{v},q^\dagger} \frac{1}{2}\|\mathscr{F}^{-1}\tilde{v} - A^\top q^\dagger\|^2 + G_{\|\cdot\|_{2,1}}(q^\dagger) + \beta\|\tilde{v}\|_1$$

e.g. via PALM*, i.e.

$$\tilde{v}^{k+1} = \mathrm{prox}_{\beta\|\cdot\|_1}\left(\tilde{v}^k - \tau\left(\tilde{v}^k - \mathscr{F}A^\top q^k\right)\right),$$

$$q^{k+1} = q^k - \sigma\left(A\left(A^\top q^k - \mathscr{F}^{-1}\tilde{v}^{k+1}\right) + \mathrm{prox}_{\|\cdot\|_{2,1}}\left(q^k + Au^\dagger\right) - Au^\dagger\right).$$

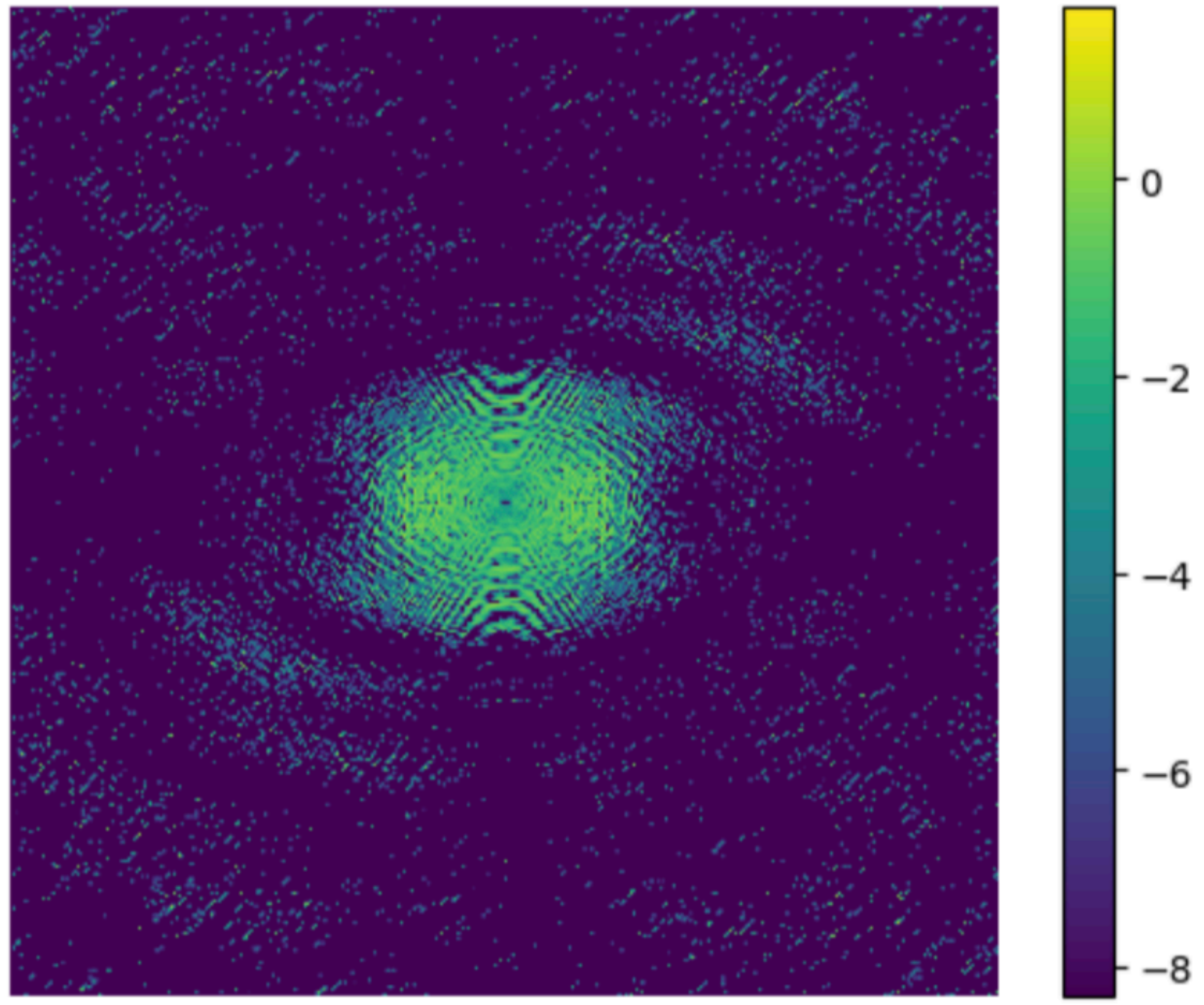
*Bolte, J., Sabach, S., & Teboulle, M. (2014). Proximal alternating linearized minimization for nonconvex and nonsmooth problems. *Mathematical Programming*, *146*(1-2), 459-494.

Numerical results

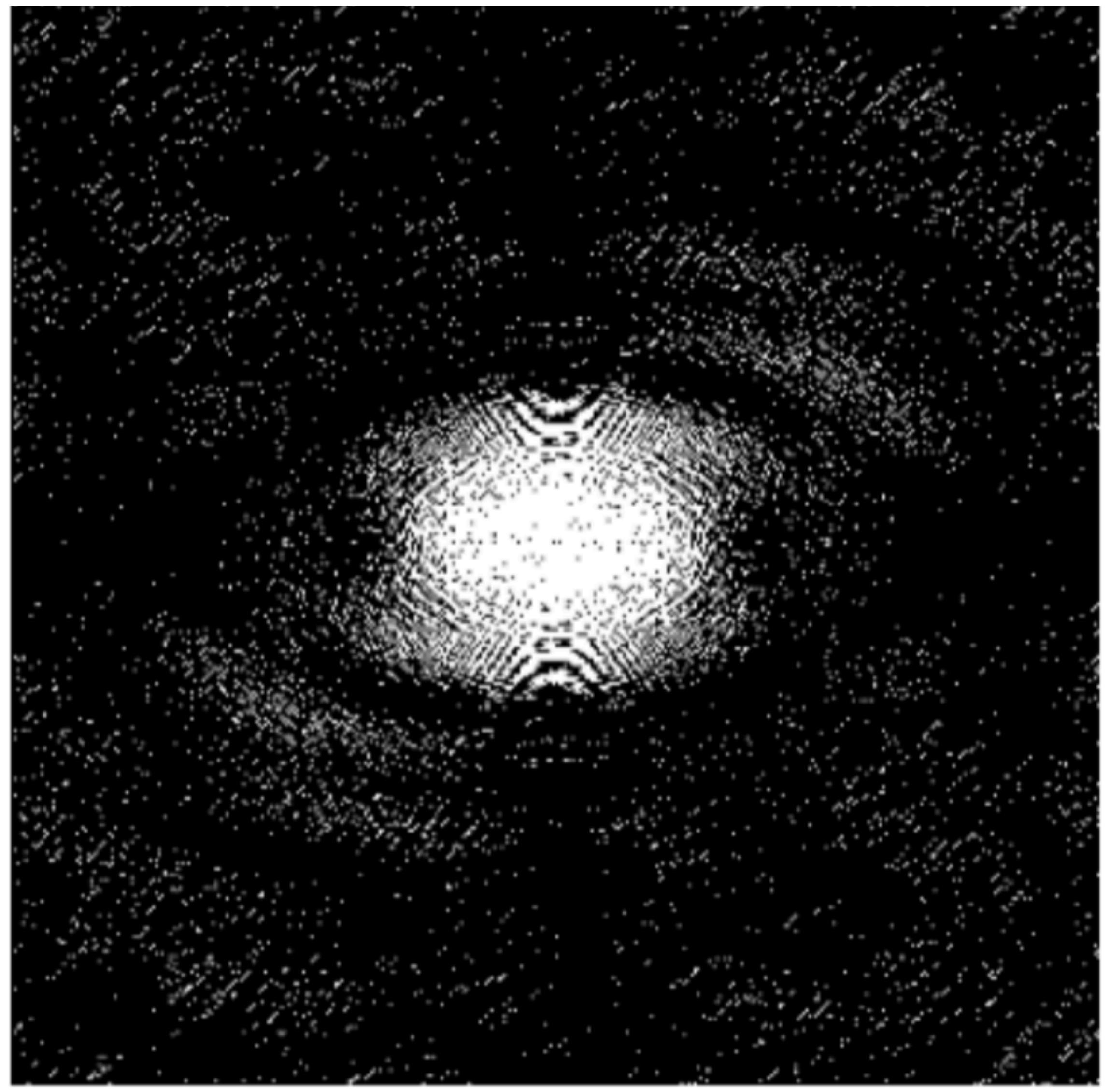
(a) Shepp-Logan phantom

(b) Eileen Collins

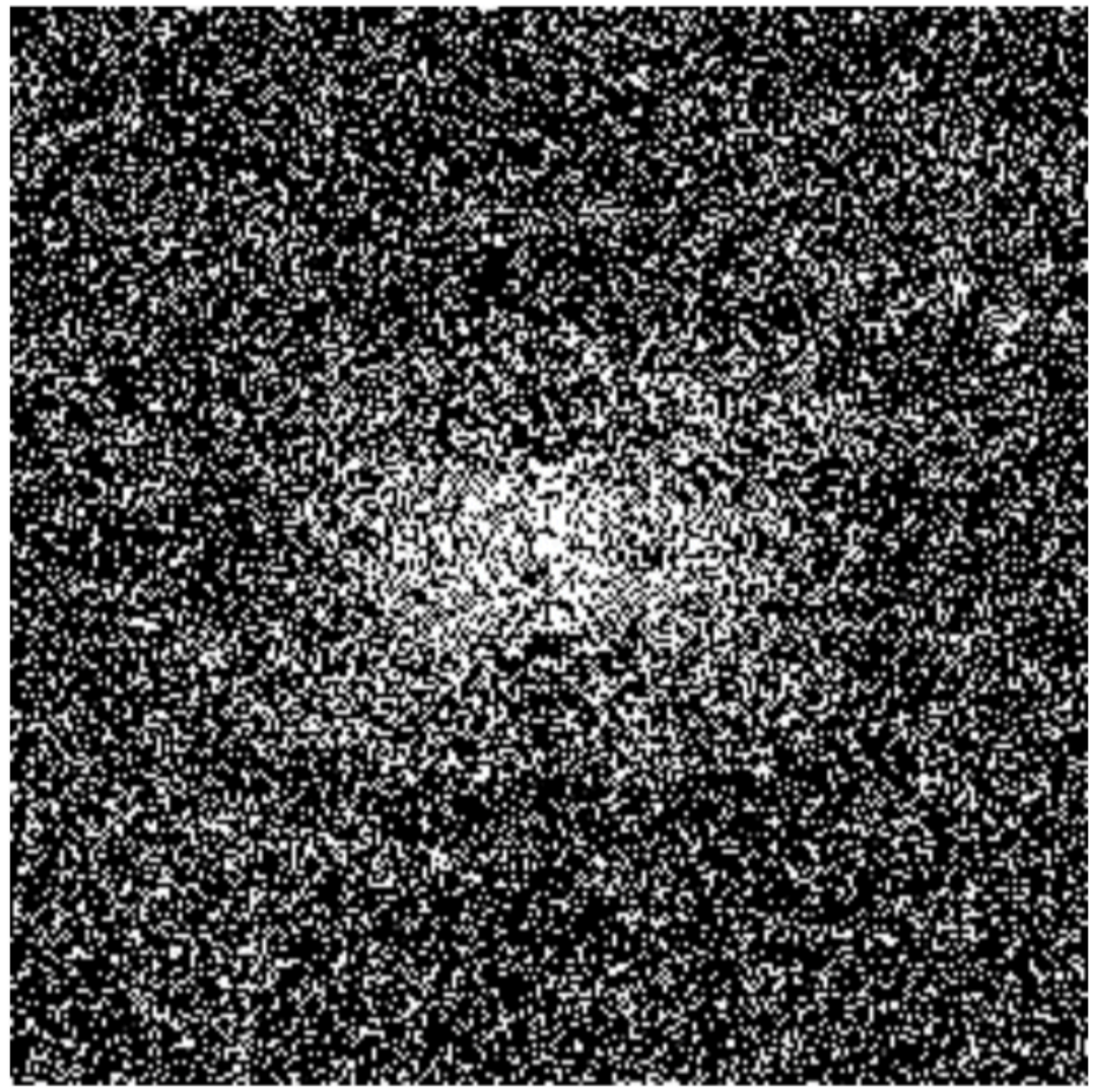
$\beta = 0.1$



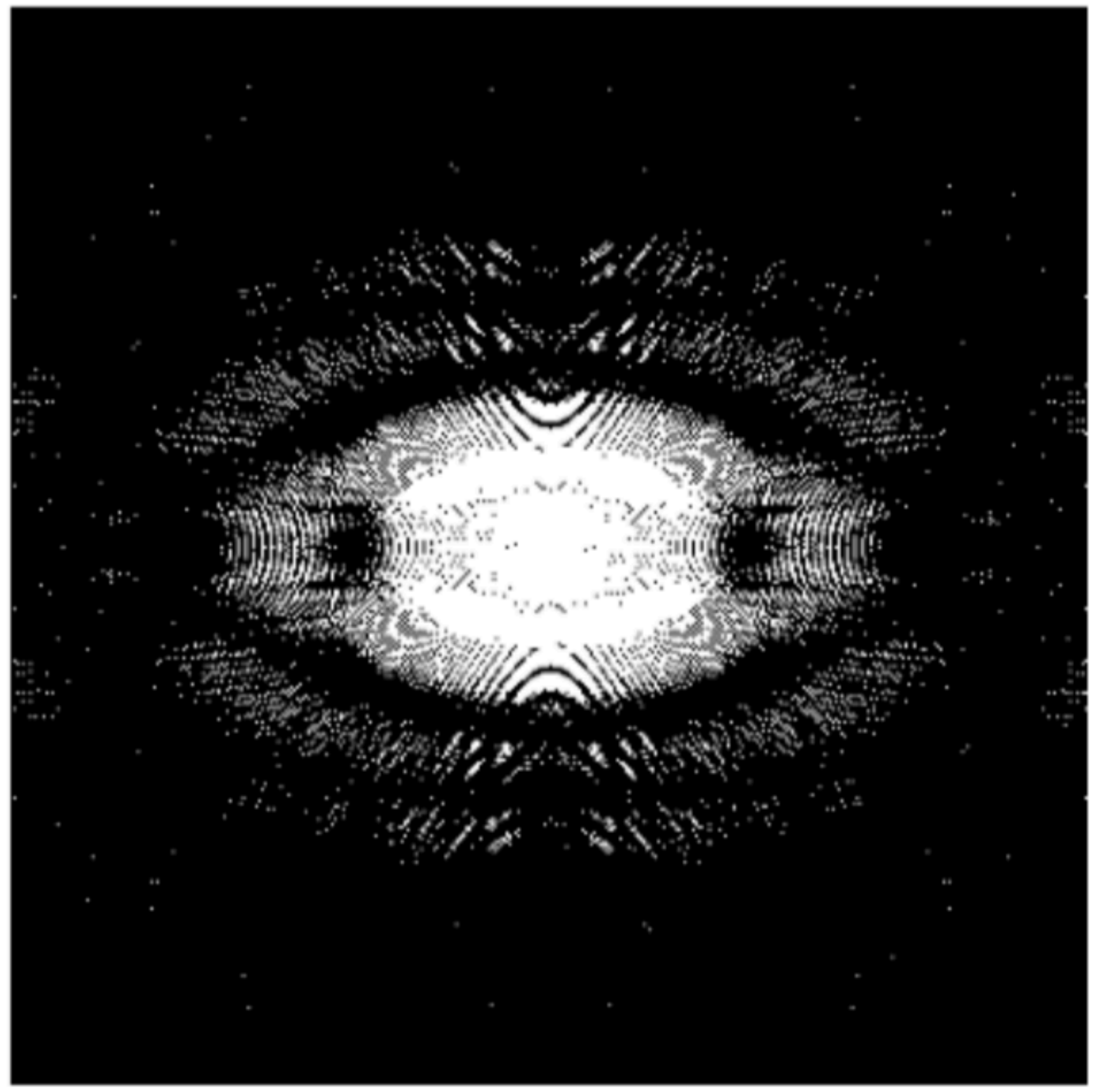
(a) $\tilde{v}^K$



(b) Sampling pattern

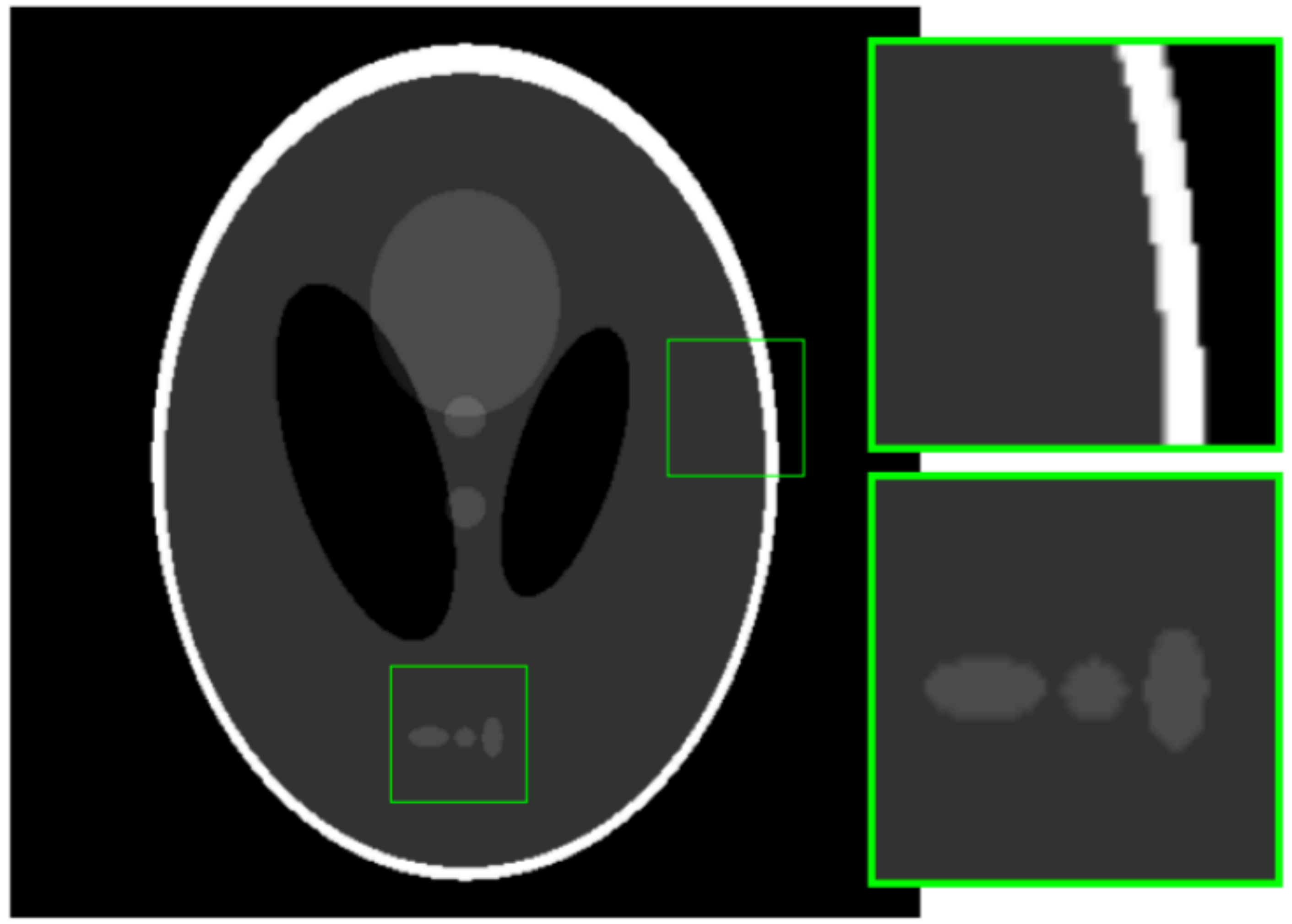


(c) Learned sampling pattern from *

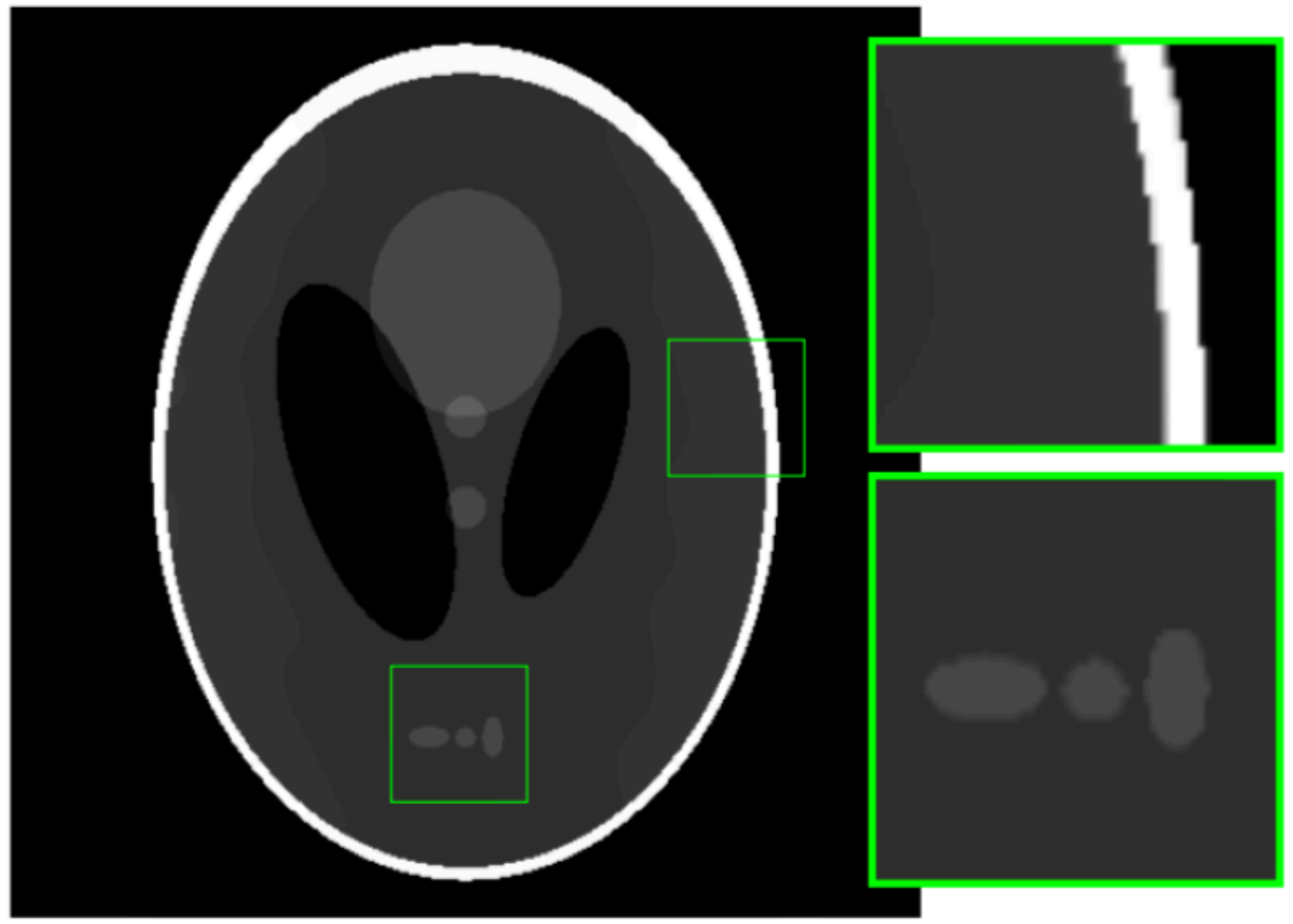


(d) Largest Fourier coefficients

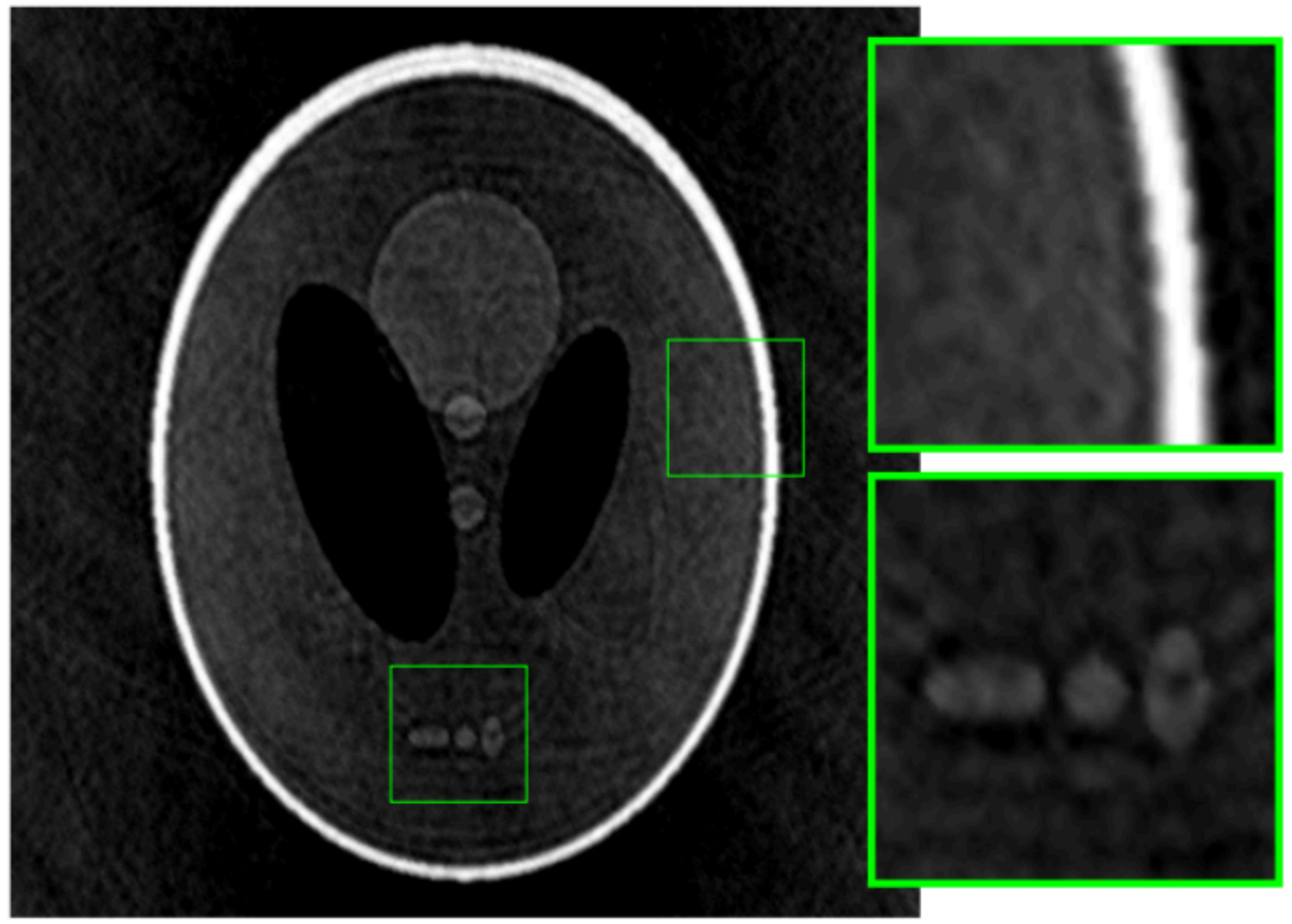
*F. Sherry et al. (2020). Learning the sampling pattern for MRI. *IEEE Transactions on Medical Imaging*, 39(12), 4310-4321.



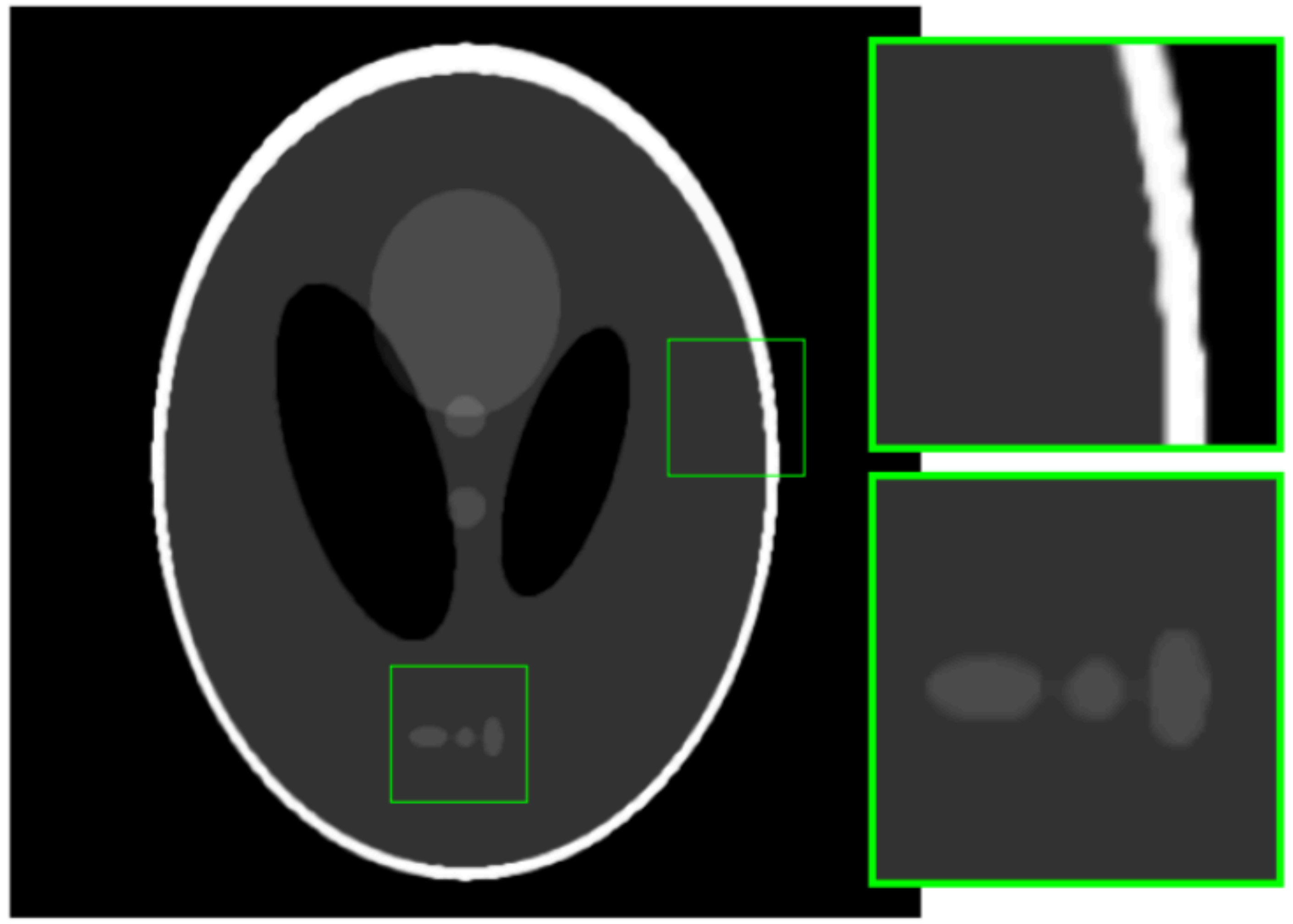
(a) ground truth $u^\dagger$



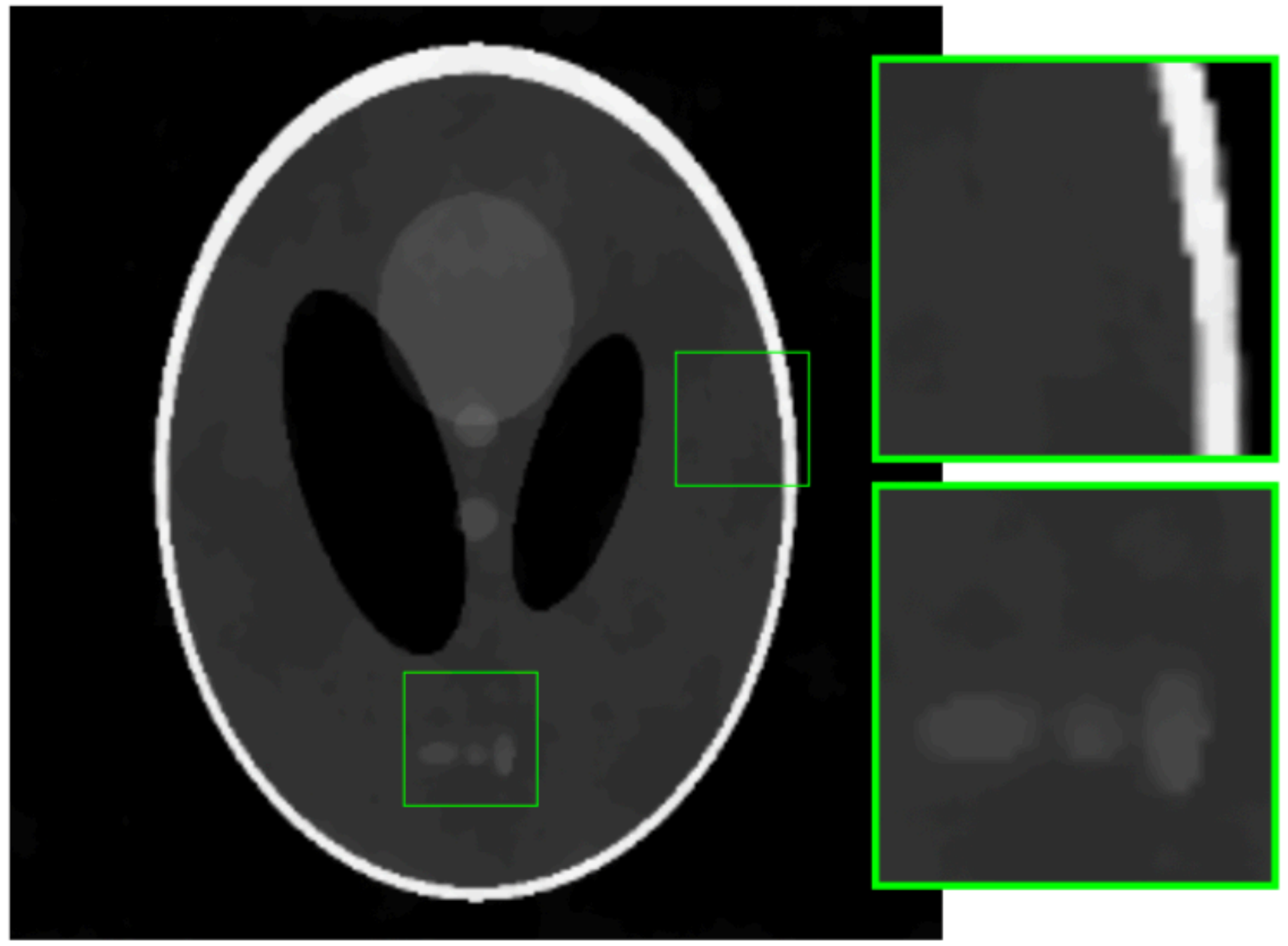
(b) approximate solution u^N



(c) projection $\mathcal{F}^{-1}S^\top S\mathcal{F}u^\dagger$



(d) low-pass filtered rec.



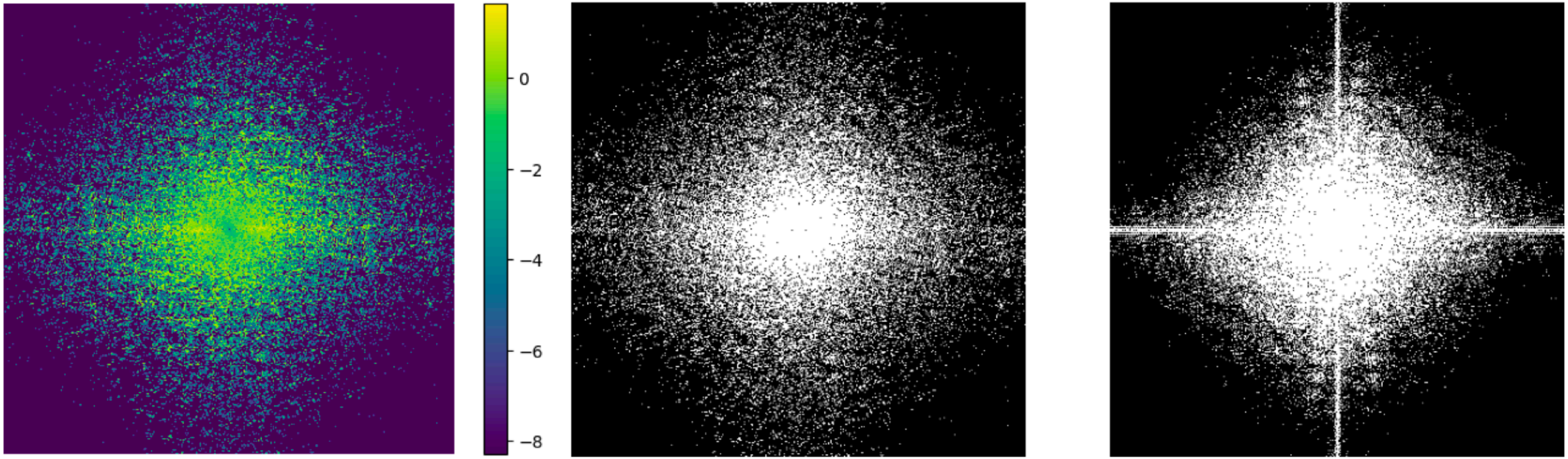
(e) reconstruction from Sherry et al.

Numerical results

(a) Shepp-Logan phantom

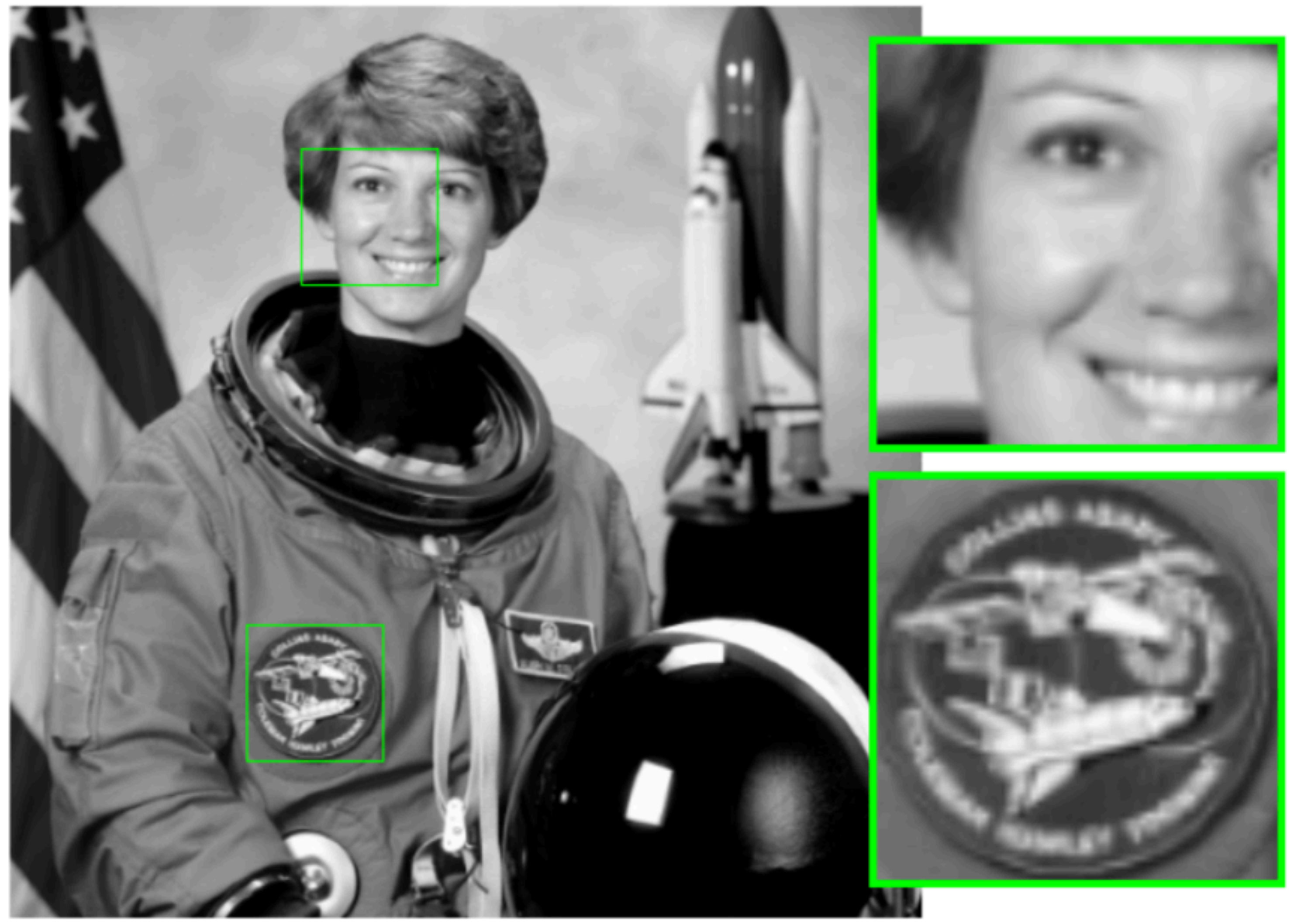
(b) Eileen Collins

Numerical results

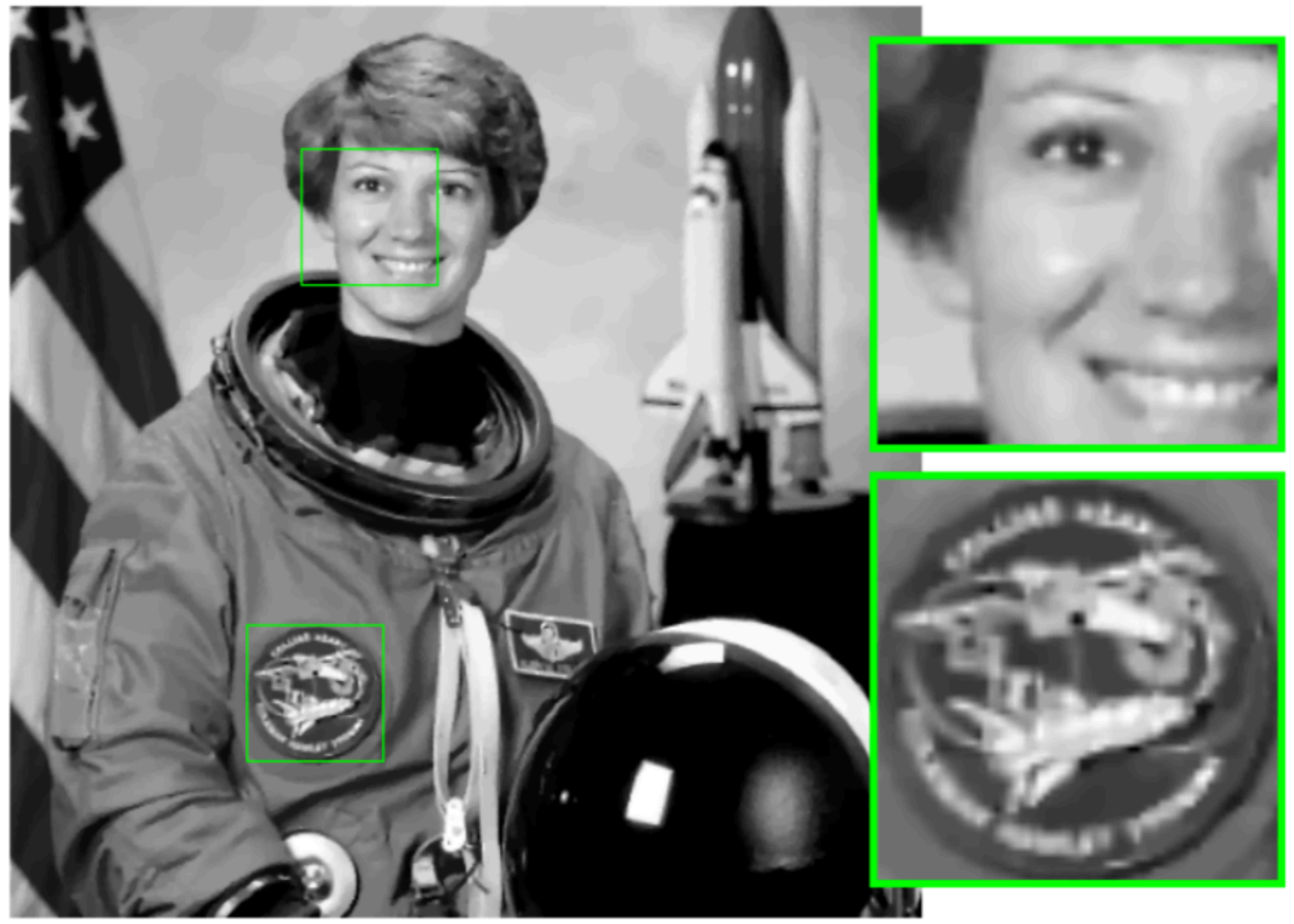


(a) $\tilde{v}^K$ (b) Sampling pattern (c) Largest Fourier coefficients

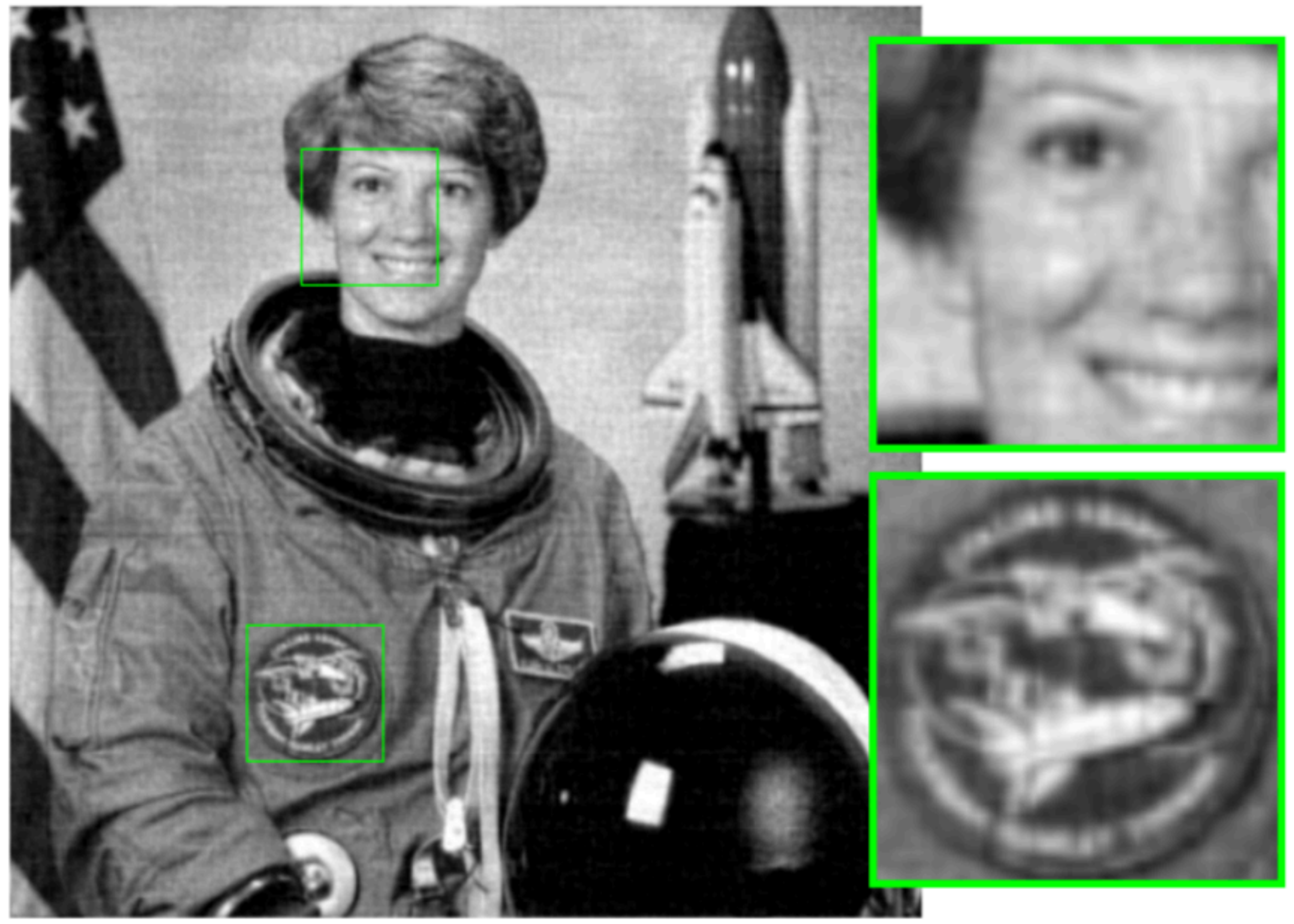
$$\beta = 0.24$$



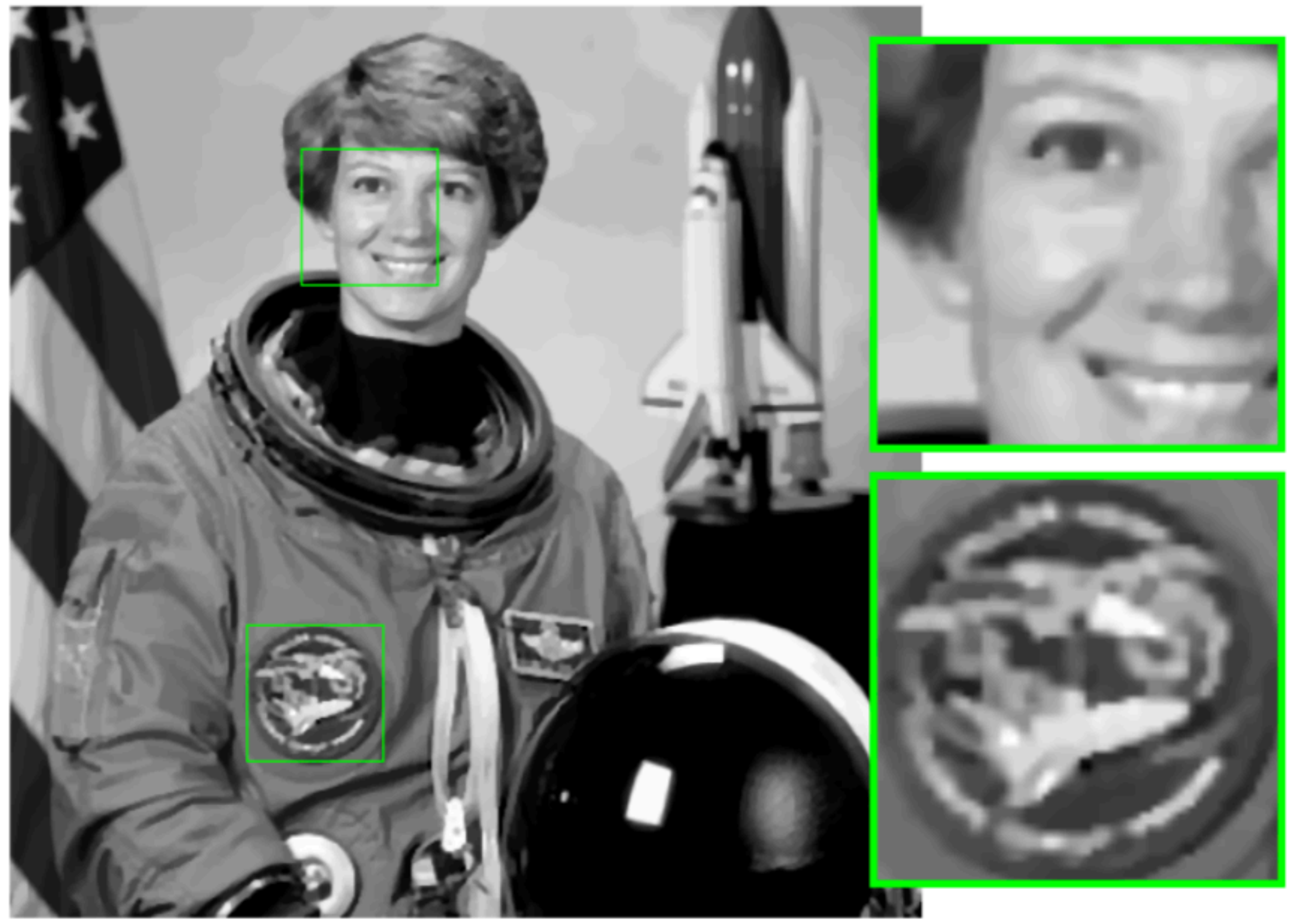
(a) ground truth $u^\dagger$



(b) approximate solution u^N



(c) projection $\mathcal{F}^{-1}S^\top S\mathcal{F}u^\dagger$



(d) low-pass filtered rec.

Conclusions & Outlook

Conclusions & outlook

Conclusions: we have

- reformulated source and range conditions as the solution of convex minimisation problems
- provided iterative algorithms for their numerical approximation
- made an attempt at supervised learning of regularisation functions with optimal error constants

Outlook:

- Higher-order or variational source conditions, gen. eigenfunctions
- Optimal sampling in various domains (e.g. MRI, single-pixel camera)
- Supervised learning for more general operator correction
- Use framework in other contexts (e.g., training deep neural networks)

Thank you for your attention!

Acknowledgements:

Isaac Newton Institute for Mathematical Sciences

Queen Mary University of London Digital Environment Research Institute

The Alan Turing Institute

EPSRC

iNδAM

IMA Journal of Applied Mathematics

hxae008

IMA Journal of Applied Mathematics (2024) **00**, 1–32
https://doi.org/10.1093/imamat/hxae008

Trust your source: quantifying source condition elements for variational regularisation methods

Error estimates in inverse problems

Example: variational regularisation

$$R_\alpha : Y \rightrightarrows X, \quad R_\alpha : f \mapsto u_\alpha \in \arg\min_{u \in X} \left\{ \frac{1}{2} \|Ku - f\|_Y^2 + \alpha J(u) \right\}$$

Assume source condition:

$$\exists\, v \in Y: \qquad K^* v \in \partial J(u^\dagger)$$

We can then show the following (well-known) estimate:

Optimality condition $u_\alpha \in R_\alpha(f^\delta)$ if and only if

$$\exists\, p_\alpha \in \partial J(u_\alpha): \qquad K^*(Ku_\alpha - f^\delta) + \alpha p_\alpha = 0$$

Burger, M., Resmerita, E., & He, L. (2007). Error estimation for Bregman iterations and inverse scale space methods in image restoration. *Computing*, *81*(2-3), 109-135.
Scherzer, O., Grasmair, M., Grossauer, H., Haltmeier, M., & Lenzen, F. (2009). Variational methods in imaging.
MB, & Burger, M. (2018). Modern regularization methods for inverse problems. *Acta Numerica*, *27*, 1-111.

Error estimates in inverse problems

$$K^*(Ku_\alpha - f^\delta) + \alpha p_\alpha - \alpha K^* v = -\alpha K^* v$$

Dual product with $u_\alpha - u^\dagger$:

$$\langle Ku_\alpha - f^\delta, Ku_\alpha - f\rangle_Y + \alpha \underbrace{\langle u_\alpha - u^\dagger, p_\alpha - K^* v\rangle_*}_{= D_J^{p_\alpha}(u^\dagger, u_\alpha) + D_J^{K^* v}(u_\alpha, u^\dagger)} = -\alpha\langle K^* v, u_\alpha - u^\dagger\rangle_*$$

where

$$D_J^p(v,u) = J(v) - J(u) - \langle p, v - u\rangle \qquad \text{for} \qquad p \in \partial J(u)$$

is the Bregman distance (divergence) w.r.t. J for arguments v, u

$D_J^{\text{symm}}(u,v) = D_J^p(v,u) + D_J^q(u,v)$ is a symmetric Bregman distance

Bregman, L. M. (1967). The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR computational mathematics and mathematical physics*, *7*(3), 200-217.
Kiwiel, K. C. (1997). Proximal minimization methods with generalized Bregman functions. *SIAM journal on control and optimization*, *35*(4), 1142-1168.

Error estimates in inverse problems

$$\langle Ku_\alpha - f^\delta, Ku_\alpha - f\rangle_Y = \frac{1}{2}\|Ku_\alpha - f\|_Y^2 + \frac{1}{2}\|Ku_\alpha - f^\delta\|_Y^2 - \frac{1}{2}\|f - f^\delta\|_Y^2$$

Hence,

$$\frac{1}{2}\|Ku_\alpha - f\|_Y^2 + \frac{1}{2}\|Ku_\alpha - f^\delta\|_Y^2 + \alpha D_J^{\text{symm}}(u_\alpha, u^\dagger) = \frac{1}{2}\|f - f^\delta\|_Y^2 - \alpha\langle v, Ku_\alpha - f\rangle_Y$$

Using the identity

$$\langle \alpha v, f - Ku_\alpha\rangle_Y = \frac{\alpha^2}{2}\|v\|_Y^2 + \frac{1}{2}\|Ku_\alpha - f\|_Y^2 - \frac{1}{2}\|\alpha v - f + Ku_\alpha\|_Y^2$$

then yields

$$\frac{1}{2}\|Ku_\alpha - f + \alpha v\|_Y^2 + \frac{1}{2}\|Ku_\alpha - f^\delta\|_Y^2 + \alpha D_J^{\text{symm}}(u_\alpha, u^\dagger) = \frac{1}{2}\|f - f^\delta\|_Y^2 + \frac{\alpha^2}{2}\|v\|_Y^2$$

Error estimates in inverse problems

$$\frac{1}{2}\|Ku_\alpha - f + \alpha v\|_Y^2 + \frac{1}{2}\|Ku_\alpha - f^\delta\|_Y^2 + \alpha D_J^{\text{symm}}(u_\alpha, u^\dagger) = \frac{1}{2}\|f - f^\delta\|_Y^2 + \frac{\alpha^2}{2}\|v\|_Y^2$$

Dividing by α and using $\|f - f^\delta\| \leq \delta$ yields the estimate

$$\frac{1}{2\alpha}\|Ku_\alpha - f + \alpha v\|_Y^2 + \frac{1}{2\alpha}\|Ku_\alpha - f^\delta\|_Y^2 + D_J^{\text{symm}}(u_\alpha, u^\dagger) \leq \frac{\alpha}{2}\|v\|_Y^2 + \frac{\delta^2}{2\alpha}$$

Hence, if we choose $\alpha(\delta) = \delta / \|v\|_Y$, we obtain

$$D_J^{\text{symm}}(u_{\alpha(\delta)}, u^\dagger) \leq \|v\|_Y \, \delta$$

How do we compute v in order to estimate (or even control) $\|v\|_Y$?