The Underwater Image Formation Model proposes that an underwater image I is obtained from the "original" image J via the following equation:

$$I(x) = J(x) \cdot t(x) + B\,(1 - t(x)) + ((J \cdot t) * g)\,(x) + n(x),$$

where t is the medium transmission, g is the point-spread function, B is the underwater background light and n is random noise.
This equation is decomposed into $I = I_1 + I_2$, where

$$I_1(x) = J(x) \cdot t(x) + B\,(1 - t(x))$$

and

$$I_2(x) = ((J \cdot t)) * g)\,(x) + n(x).$$

I_1 is first estimated via analytic means, which also yield an estimation for t. Then, I_2 is taken as $I_2 = J - I_1$. Finally J and g are estimated variationally. Define

$$E(J, g) = \|g * J - I_2\|_2^2 + \mathcal{R}_1(J) + \mathcal{R}_2(g),$$

where $\mathcal{R}_1$ and $\mathcal{R}_2$ are regularization functions. Thus, J and g are obtained by solving the following minimization problem:

$$J, g = \operatorname{argmin}_{\hat{J},\hat{g}} E(\hat{J}, \hat{g}).$$

Given the complexity of such problem, a greedy approach is used, whereby E is alternatively minimized with respect to one of the variables while fixing the other. Thus, the proposed algorithm is the following:

$$\begin{cases} J_{n+1} = \operatorname*{argmin}_{\hat{J}} E(\hat{J}, \hat{g}_n) & (1) \\ g_{n+1} = \operatorname*{argmin}_{\hat{g}} E(\hat{J}_{n+1}, \hat{g}) & (2) \end{cases}$$

Each of the iteration steps above presents a minimization problem of its own. Due to the symmetrical nature of E, it is possible to solve both problems with the same approach, which we describe in the following.
Let Ω be an open, finitely-measured subset of $\mathbb{R}^n$ and denote by $L^2(\Omega)$ the Hilbert space formed by the square-integrable real-valued functions on Ω, with the usual identification of almost-everywhere-identical functions. Let $y, s \in L^2(\Omega)$ with $y \neq 0$, and $R\colon L^2(\Omega) \to \mathbb{R}$ convex. We shall refer to s as "shift" and to R as "regularitazion function". Consider the proper and strictly convex functional

$$F(x; y, s, R) = \frac{1}{2}\|x * y - s\|^2 + R(x).$$

We aim to solve the following optimization problem, which by the direct method admits a unique solution:

$$\operatorname*{argmin}_{x} F(x; y, s, R) \tag{3}$$

The Chambolle-Pock primal-dual algorithm can be applied in a standard fashion to yield iterates for this problem. In order to apply it, it is necessary to introduce some additional notation.

Consider the endomorphism $K \in \mathcal{L}\left(L^2(\Omega), L^2(\Omega)\right)$ given by $K(x) = x * y$. To see that this is, in fact, an endomorphism, we employ the fact that, since Ω is finitely-measured, then $L^2(\Omega) \subseteq L^1(\Omega)$, together with the Young inequality: since $y \in L^2(\Omega) \subseteq L^1(\Omega)$, for all $x \in L^2(\Omega)$ it follows that

$$\|x * y\|_2 \leq \|x\|_2 \cdot \|y\|_1 < +\infty.$$

Thus $x * y \in L^2(\Omega)$. This inequality also implies the boundedness of K. Linearity is immediate.
Now define $N\colon L^2(\Omega) \to \mathbb{R}$ by $N(a) = \frac{1}{2}\|a - s\|_2^2$. It is clear that N is strictly convex. Thus, we can re-express our functional in the form

$$F(x; y, s, R) = N(K(x; y); s) + R(x), \tag{4}$$

with which the Chambolle-Pock algorithm has its hypotheses satisfied and can be applied. It yields the following iterates:

$$\begin{cases} x_{n+1} = \text{prox}_{\tau R}(x_n - \tau K^* z_n) \\ \overline{x}_{n+1} = 2x_{n+1} - x_n \\ z_{n+1} = \text{prox}_{\sigma N^*}(z_n + \sigma K \overline{x}_{n+1}) \end{cases} \tag{5}$$

Here, it is possible to obtain analytical expressions for both $\text{prox}_{\sigma N^*}$ and K^*. They are given by

$$\text{prox}_{\sigma N^*}(x) = \frac{x - \sigma s}{\sigma + 1} \quad \text{and} \quad K^* z = z * \overline{y}, \tag{6}$$

with $\overline{y} = y \circ (-Id)$. Detailed deductions for these expressions can be found in section 1.
We propose to unfold $\text{prox}_{\tau R}$; that is, substitute it by a neural network.

1 Computations

1.1 Proximity operator

Let $N\colon L^2(\Omega) \to \mathbb{R}$ be as in eq. (4). In this subsection, the deduction of an analytical expression of the proximity operator $\text{prox}_{\sigma N^*}$ will be presented.
By the Moreau decomposition theorem,

$$\text{prox}_{\sigma N^*}(x) = x - \sigma \text{prox}_{N/\sigma}\left(\frac{x}{\sigma}\right). \tag{7}$$

Thus, it is sufficient to compute $\text{prox}_{\lambda N}$ for an arbitrary $\lambda > 0$ and then replace $\lambda = \frac{1}{\sigma}$.
For a fixed $z \in L^2(\Omega)$, consider the functional $J_{z,\lambda}(y) = \frac{\|y-z\|_2^2}{2\lambda} + \frac{1}{2}\|y - s\|_2^2$, so that

$$\text{prox}_{\lambda N}(z) = \text{argmin}_y J_{z,\lambda}(y).$$

Clearly, $J_{z,\lambda}$ is Gâteaux differentiable and

$$J_{z,\lambda}(y; h) = \langle h, \frac{1}{\lambda}(y - z)\rangle + \langle h, y - s\rangle.$$

Whence $\forall h\colon J_{z,\lambda}(y; h) = 0$ if, and only if, $y = \frac{z+\lambda s}{\lambda+1}$. Thus, by the Fermat principle, and taking $\lambda = \frac{1}{\sigma}$,

$$\text{prox}_{N/\sigma}(z) = \text{argmin}_y J_{z,1/\sigma}(y) = \frac{\sigma z + s}{\sigma + 1}.$$

Combining this last equation with **??**, we have that

$$\operatorname{prox}_{\sigma N^*}(x) = x - \frac{\sigma}{\sigma+1}(x+s) = \frac{x-\sigma s}{\sigma+1} \tag{8}$$

1.2 Convolution

Let $K\colon L^2(\Omega) \to L^2(\Omega)$ be as in eq. (4). In this subsection, the deduction of an analytical expression of the adjoint operator $K^*\colon L^2(\Omega) \to L^2(\Omega)$ will be presented.
By definition, K^* is the only operator satisfying the equality

$$\langle Ka, b\rangle = \langle a, K^*b\rangle$$

for all $a, b \in L^2(\Omega)$, where $\langle\cdot,\cdot\rangle$ is the usual scalar product in $L^2(\Omega)$: $\langle f, g\rangle = \int_\Omega fg\ dm$, where m is the Lebesgue measure restricted to Ω.
Note that, in order to properly define convolution between functions with domain $\Omega \subseteq \mathbb{R}^n$, it is first necessary to extend to $\mathbb{R}^n$ by setting them to 0 in Ω^c.
It is well-known that for any given $\varphi, \phi, \psi\colon \mathbb{R}^n \to \mathbb{R}$ such that their pairwise convolutions are defined and integrable, the following equality holds:

$$\langle \varphi * \phi, \psi\rangle = \langle \phi, \psi * \overline{\varphi}\rangle \tag{9}$$

where $\overline{\varphi}(x) = \varphi(-x)$. Therefore,

$$\langle a * y, b\rangle = \langle a, b * \overline{y}\rangle.$$

Thus, K^* is simply given by $K^*b = b * \overline{y}$.

2 On the discrete convolution.

Definition 1. *Define a **centered image domain** as a subset $D \subseteq \mathbb{R}$ of the form $D = [d,u] \times [l,r] \cap \mathbb{Z}^2$, where $d,u,l,r \in \mathbb{Z}$, $d \leq 0 \leq u$ $l \leq 0 \leq r$, and both $-1 \leq d+u \leq 0$ and $-1 \leq l+r \leq 0$.*

This means taking the usual representation for the domains of (grayscale) digital images (from 0 to the height or the width minus one), and centering it around the origin, with the negative parts of each dimension being one pixel larger than the strictly positive parts if the number of pixels in that dimension is even, and equal in size if it is odd.
Consider a grayscale image I represented as a 2D centered function taking real values:

$$I\colon D(I) =\to \mathbb{R},$$

where $D(I)$ is the unique centered image domain determined by the resolution of I. Namely, if I has a height of H and a width of W (in pixels), then

$$D(I) = \left[-\left\lceil\frac{H-1}{2}\right\rceil, \left\lfloor\frac{H-1}{2}\right\rfloor\right] \times \left[-\left\lceil\frac{W-1}{2}\right\rceil, \left\lfloor\frac{W-1}{2}\right\rfloor\right] \tag{10}$$

Conversely, if an image I has a centered domain $D(I) = [d,u] \times [l,r] \cap \mathbb{Z}^2$, its height H and width W (in total pixels) are given by

$$H = u - d + 1, \qquad W = r - l + 1 \tag{11}$$

It is clear that one may extend this representation to an arbitrarily larger domain using different methods. This is usually referred to as **padding**. For instance, one could add zeros or mirror I around its edges forming an infinite pattern. It is also possible to do the converse: going from a given domain to an arbitrarily smaller domain by simply restricting it.

Definition 2. *Given a* ***kernel*** g *and an* ***image*** I*, we define the* ***convolution*** $g * I$ *in the domain* D *as follows: for each* $x \in D$*,*

$$g * I(x) = \sum_{y \in D(g)} g(y)I(x-y),$$

where I *has been extended if necessary to accommodate for out-of-bounds values* $x - y$*.*

Note how one may **choose** the domain of $g * I$ by simply padding I as necessary. However, the case where I needs not be extended and the domain is maximal is of particular interest. It is simple to check that this happens whenever $x = (x_1, x_2)$ satisfies the constraints

$$\begin{array}{c} d_I + u_g \leq x_1 \leq d_g + u_I \\ l_I + r_g \leq x_2 \leq l_g + r_I \end{array},$$

meaning that the output image would have height $(u_I - u_g) - (u_g - d_g) + 1 = H_I - H_g + 1$ and width $(r_I - l_I) - (r_g - l_g) + 1 = W_I - W_g + 1$, where H_I, W_I and H_g, W_g are the height and width (in pixels) of I and g, respectively.
If one wishes to obtain a domain $D = [d, u] \cap [l, r] \cap \mathbb{Z}^2$ with height $H \geq H_I - W_g + 1$ and width $W \geq W_I - W_g + 1$, it is possible to characterize the padding needed in I: call d_p, u_p, l_p and r_p the bottom, top, left and right padding required, respectively (all are nonnegative quantities). Then, the following equalities should be satisfied:

$$\begin{cases} d_I - d_p = d - u_g \\ u_I + u_p = u - d_g \\ l_I - l_p = l - r_g \\ r_I + r_p = r - l_g \end{cases},$$

from which it is trivial to obtain the padding amounts. Note that this may be combined with eq. (10) in order to obtain the padding amounts from the (more intuitive) height and width of the desired output.

Theorem 1. *For a fixed domain* D*, zero-padded convolution is commutative:*

$$g * ID = I * gD.$$

Proof. For any given $x \in D$, since I is being zero-padded,

$$g * I(x) = \sum_{y \in D(g) \cap (x - D(I))} g(y)I(x-y),$$

where $x - D(I)$ is the usual set addition notation $b + A = \{b + a | a \in A\}$. Now introduce a variable change $z = x - y$ (invertible via $y = x - z$) and note that $y \in D(g) \cap (x - D(I))$ if, and only if, $z \in (x - D(g)) \cap D(I)$. Therefore,

$$\sum_{y \in D(g) \cap (x - D(I))} g(y)I(x-y) = \sum_{z \in (x - D(g)) \cap D(I)} g(x-z)I(z),$$

and the last term is clearly $I * g(x)$. □

2.1 Derivative of convolution.

Consider the following functional:

$$\tilde{F}(g) = \frac{1}{2}\|g * J - I_2\|_2^2,$$

where J and I_2 are images of the same resolution and the convolution domain is chosen so as to have the same resolution as them.

Claim: $\tilde{F}$ is Fréchet differentiable and $\nabla F(g) = (g * J - I_2) * \overline{J}$, where the inner convolution has the same domain as I_2, the outer convolution has the same domain as g, and $\overline{J}$ is given by $\overline{J}(x) = J(-x)$.

Proof. First note that, for a given $k \in D(g)$

$$\frac{\partial (g * J)}{\partial g(k)} = \frac{\partial \left(\sum_{y \in D(g)} g(y) J(x-y)\right)}{\partial g(k)} = J(x-k).$$

Now, fully "unroll" $\tilde{F}$ into a summatory

$$\tilde{F}(g) = \frac{1}{2} \sum_{x \in D(J)} \left(g * J(x) - I_2(x)\right)^2.$$

Combining the two previous equalities and the chain rule, we obtain

$$\frac{\partial \tilde{F}}{\partial g(k)} = \sum_{x \in D(J)} (g * J(x) - I_2(x)) \cdot J(x-k) = \sum_{x \in D(J)} \big(g * J - I_2)(x) \cdot \overline{J}(k-x) = \big((g * J - I_2) * \overline{J}\big)(k).$$

Therefore, $\nabla \tilde{F}(g) = \big(g * J - I_2\big) * \overline{J}$, as claimed. □