

This is the final peer-reviewed accepted manuscript of:

Pragliola, M., Calatroni, L., Lanza, A., Sgallari, F. (2021). Residual Whiteness Principle for Automatic Parameter Selection in ℓ_2 - ℓ_2 Image Super-Resolution Problems. In: Elmoataz, A., Fadili, J., Quéau, Y., Rabin, J., Simon, L. (eds) Scale Space and Variational Methods in Computer Vision. SSVM 2021. Lecture Notes in Computer Science(), vol 12679. Springer, Cham

The final published version is available online at https://dx.doi.org/10.1007/978-3-030-75549-2_38

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)

When citing, please refer to the published version.

Residual whiteness principle for automatic parameter selection in ℓ_2 - ℓ_2 image super-resolution problems

Monica Pragliola^{*1}, Luca Calatroni^{†2}, Alessandro Lanza^{‡1}, and
Fiorella Sgallari^{§1}

¹Department of Mathematics, University of Bologna, Italy
²CNRS, UCA, INRIA, Morpheme, I3S, Sophia-Antipolis, France

Abstract

We propose an automatic parameter selection strategy for variational image super-resolution of blurred and down-sampled images corrupted by additive white Gaussian noise (AWGN) with unknown standard deviation. By exploiting particular properties of the operators describing the problem in the frequency domain, our strategy selects the optimal parameter as the one optimising a suitable residual whiteness measure. Numerical tests show the effectiveness of the proposed strategy for generalised ℓ_2 - ℓ_2 Tikhonov problems.

1 Introduction

The problem of single-image Super-Resolution (SR) consists in finding a high-resolution (HR) image starting from low-resolution (LR) blurred and noisy data. The huge number of applications which benefits from the recovery of HR information, ranging from remote sensing to biomedical imaging, motivates the large amount of research still ongoing in this field.

Mathematically, the problem can be described as follows. Let $\mathbf{X} \in \mathbb{R}^{N_r \times N_c}$ denote the original HR image, with $\mathbf{x} = \text{vec}(\mathbf{X}) \in \mathbb{R}^N$, $N = N_r N_c$, being its vectorisation. The process describing the mapping from HR to LR data can be described by the following linear observation model

$$\mathbf{b} = \mathbf{S}\mathbf{K}\mathbf{x} + \mathbf{e}, \quad \text{with } \mathbf{e} \text{ realisation of } \mathbf{E} \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_n), \quad (1)$$

^{*}monica.pragliola2@unibo.it

[†]calatroni@i3s.unice.fr

[‡]alessandro.lanza2@unibo.it

[§]fiorella.sgallari@unibo.it

where $\mathbf{b}, \mathbf{e} \in \mathbb{R}^n$, $n = n_r n_c$, are the vectorised observed and noise image, respectively, both consisting of $n_r \times n_c$ pixels, $\mathbf{S} \in \mathbb{R}^{n \times N}$ is the down-sampling operator inducing a pixel decimation with factor d_r and d_c along the rows and the columns of \mathbf{X} , respectively - i.e., $N_r = n_r d_r$, $N_c = n_c d_c$ - $\mathbf{K} \in \mathbb{R}^{N \times N}$ represents a space-invariant blurring operator, $\mathbf{I}_n \in \mathbb{R}^{n \times n}$ denotes the n -dimensional identity matrix and \mathbf{E} is an n -variate Gaussian-distributed random vector with zero mean and scalar covariance matrix, with σ indicating the (unknown) noise standard deviation. We set $d = d_r d_c$, so that $N = nd$.

To overcome the ill-posedness of problem (1), one can seek an estimate \mathbf{x}^* of \mathbf{x} by minimising a suitable cost function $\mathcal{J} : \mathbb{R}^N \rightarrow \mathbb{R}^+$. In this work, we consider in particular a generalised ℓ_2 - ℓ_2 Tikhonov-regularised model of the form

$$\mathbf{x}^*(\mu) = \arg \min_{\mathbf{x} \in \mathbb{R}^N} \left\{ \mathcal{J}(\mathbf{x}; \mu) := \frac{\mu}{2} \|\mathbf{SKx} - \mathbf{b}\|_2^2 + \frac{1}{2} \|\mathbf{Lx} - \mathbf{v}\|_2^2 \right\}, \quad (2)$$

where the operator $\mathbf{L} \in \mathbb{R}^{M \times N}$ and the vector $\mathbf{v} \in \mathbb{R}^M$ are known. The data term $\|\mathbf{SKx} - \mathbf{b}\|_2^2$ encodes the AWGN assumption on \mathbf{e} , while the regularisation term $\|\mathbf{Lx} - \mathbf{v}\|_2^2$ encodes prior information on the unknown target. Finally, the *regularisation parameter* $\mu \in \mathbb{R}_*^+$ in (2) balances the action of the fidelity against regularisation; its choice is of crucial importance for high quality reconstructions.

When $\mathbf{S} = \mathbf{I}_N$, under general assumptions - see (A3)-(A4) in Sec. 1.1 - the problem in (2) can be solved very efficiently. However, the presence of a non-trivial \mathbf{S} makes the computation of the least-squares solution very costly. In [12], upon a specific choice of \mathbf{S} , the authors proposed an efficient strategy for the solution of (2), for which Generalised Cross Validation [3] is used to select the optimal μ . This is known to be impractical for large-scale problems [2].

A popular strategy which aims at overcoming the downsides of empirical parameter selection rules while exploiting the information available on the noise corruption is the celebrated discrepancy principle (DP) (see [4, 1] for general problems and [9] for applications to super-resolution problems), which can be formulated as follows:

$$\text{select } \mu = \mu^* \text{ such that } \|\mathbf{r}^*(\mu^*)\|_2 = \|\mathbf{SKx}^*(\mu) - \mathbf{b}\|_2 = \tau \sqrt{n} \sigma, \quad (3)$$

with $\mathbf{x}^*(\mu)$ being the solution of (2) and τ denoting the discrepancy coefficient. When σ is known, τ is set equal to 1, otherwise a value slightly greater than 1 is typically chosen to avoid noise under-estimation. Clearly, in real world applications an accurate estimate of σ is not available, which often limits the applicability of DP strategies.

Recently, in the context of image restoration problems, a number of works has focused on the design of variational models explicitly exploiting in their formulations the assumed whiteness of the corrupting noise - see, e.g., [5, 7]. Based on these promising results, in [6], the authors propose a strategy named *residual whiteness principle* (RWP), that relies on the whiteness property of the noise to properly set the regularisation parameter μ . The RWP automatically selects a value for μ that maximises the whiteness of the residual image

$\mathbf{r}^*(\mu) = \mathbf{S}\mathbf{K}\mathbf{x}^*(\mu) - \mathbf{b}$, or equivalently minimises the squared Euclidean norm of the normalised auto-correlation of $\mathbf{r}^*(\mu)$. The RWP has there been applied to the automatic selection of μ in Tikhonov-regularised least squares problems which are frequently encountered in iterative *alternating direction method of multipliers* (ADMM) optimisation frameworks when used to larger classes of non-smooth regularisation models.

In this paper, we extend the results obtained in [6] to SR problems of the form (2). As in [6], the proposed strategy can be easily extended to models more general than the one in (2).

1.1 Notations, preliminaries and assumptions

In the following, for $c \in \mathbb{C}$ we use \bar{c} , $|c|$ to indicate the conjugate and the modulus of c , respectively. We denote by \mathbf{F} , \mathbf{F}^H the 2D Fourier transform and its inverse, respectively. For any $\mathbf{v} \in \mathbb{R}^N$ and any $\mathbf{A} \in \mathbb{R}^{N \times N}$, we use the notations $\tilde{\mathbf{v}} = \mathbf{F}\mathbf{v}$ and $\tilde{\mathbf{A}} = \mathbf{F}\mathbf{A}\mathbf{F}^H$ to denote the action of the 2D Fourier transform operator \mathbf{F} on vectors and matrices, respectively. Given a permutation matrix $\mathbf{P} \in \mathbb{R}^{N \times N}$, we denote by $\hat{\mathbf{v}} = \mathbf{P}\tilde{\mathbf{v}}$ and by $\hat{\mathbf{A}} = \mathbf{P}\tilde{\mathbf{A}}\mathbf{P}^T$ the action of \mathbf{P} on the Fourier-transformed vector $\tilde{\mathbf{v}}$ and matrix $\tilde{\mathbf{A}}$, respectively. Finally, by $\check{\mathbf{A}}$ we denote the product $\check{\mathbf{A}} = \mathbf{P}\tilde{\mathbf{A}}^H\mathbf{P}^T$, i.e. the action of \mathbf{P} on $\tilde{\mathbf{A}}^H$.

We recall some results that will be useful in the following discussion and a well-known property of the Kronecker product ‘ \otimes ’.

Lemma 1 ([10]). *Let $\mathbf{J}_d \in \mathbb{R}^{d \times d}$ denote a matrix of ones. We have:*

$$\widetilde{\mathbf{S}^H\mathbf{S}} = \frac{1}{d}(\mathbf{J}_{d_r} \otimes \mathbf{I}_{n_r}) \otimes (\mathbf{J}_{d_c} \otimes \mathbf{I}_{n_c}). \quad (4)$$

Lemma 2. *Let $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ be given matrices such that \mathbf{AC}, \mathbf{BD} exist. We have:*

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC} \otimes \mathbf{BD}). \quad (5)$$

Lemma 3 (Woodbury formula). *Let $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4$ matrices and let \mathbf{A}_1 and \mathbf{A}_3 be invertible. Then, the following inversion formula holds:*

$$(\mathbf{A}_1 + \mathbf{A}_2\mathbf{A}_3\mathbf{A}_4)^{-1} = \mathbf{A}_1^{-1} + \mathbf{A}_1^{-1}\mathbf{A}_2(\mathbf{A}_3^{-1} + \mathbf{A}_4\mathbf{A}_1^{-1}\mathbf{A}_2)^{-1}\mathbf{A}_4\mathbf{A}_1^{-1}. \quad (6)$$

The results recalled and proposed in this paper rely on the following assumptions on the image formation model and on the linear operators $\mathbf{S}, \mathbf{K}, \mathbf{L}$.

- (A1) The original image \mathbf{X} is assumed to be square, i.e. $N_r = N_c$, and $d_c = d_r$.
- (A2) The conjugate transpose $\mathbf{S}^H \in \mathbb{R}^{N \times n}$ of the down-sampling operator interpolates the decimated image with zeros, and $\mathbf{S}\mathbf{S}^H = \mathbf{I}_n$.
- (A3) The matrices \mathbf{S}, \mathbf{K} and \mathbf{L} in (2) are such that $\text{null}(\mathbf{SK}) \cap \text{null}(\mathbf{L}) = \mathbf{0}_N$, with $\mathbf{0}_N$ denoting the N -dimensional null vector.

(A4) As a consequence of the space-invariance of the blur, the matrix \mathbf{K} represents a 2D discrete convolution operator. Also the regularisation matrix \mathbf{L} is required to represent a 2D convolutional operator, so that \mathbf{K} and \mathbf{L} can be diagonalised by the 2D discrete Fourier transform. In formula:

$$\mathbf{K} = \mathbf{F}^H \mathbf{\Lambda} \mathbf{F} \quad \text{and} \quad \mathbf{L} = \mathbf{F}^H \mathbf{\Gamma} \mathbf{F}, \quad \text{with} \quad \mathbf{F}^H \mathbf{F} = \mathbf{F} \mathbf{F}^H = \mathbf{I}_N, \quad (7)$$

where $\mathbf{\Lambda}, \mathbf{\Gamma} \in \mathbb{C}^{N \times N}$ are diagonal matrices defined by

$$\mathbf{\Lambda} = \text{diag}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_N), \quad \mathbf{\Gamma} = \text{diag}(\tilde{\gamma}_1, \dots, \tilde{\gamma}_N).$$

Notice that assumption (A3) guarantees the existence of global minimisers for the cost function $\mathcal{J}(\cdot; \mu) : \mathbb{R}^N \rightarrow \mathbb{R}^+$ in (2).

2 Residual whiteness principle

Let us consider the noise realisation \mathbf{e} in (1) in its original $n_r \times n_c$ matrix form:

$$\mathbf{e} = \{e_{i,j}\}_{(i,j) \in \Omega}, \quad \Omega := \{0, \dots, n_r - 1\} \times \{0, \dots, n_c - 1\}.$$

The *sample auto-correlation* $a : \mathbb{R}^{n_r \times n_c} \rightarrow \mathbb{R}^{(2n_r - 1) \times (2n_c - 1)}$ of realisation \mathbf{e} is

$$a(\mathbf{e}) = \{a_{l,m}(\mathbf{e})\}_{(l,m) \in \Theta}, \quad \Theta := \{-(n_r - 1), \dots, n_r - 1\} \times \{-(n_c - 1), \dots, n_c - 1\},$$

with each scalar component $a_{l,m}(\mathbf{e}) : \mathbb{R}^{n_r \times n_c} \rightarrow \mathbb{R}$ given by

$$a_{l,m}(\mathbf{e}) = \frac{1}{n} (\mathbf{e} \star \mathbf{e})_{l,m} = \frac{1}{n} (\mathbf{e} \ast \mathbf{e}')_{l,m} = \frac{1}{n} \sum_{(i,j) \in \Omega} e_{i,j} e_{i+l,j+m}, \quad (l,m) \in \Theta, \quad (8)$$

where index pairs (l,m) are commonly called *lags*, \star and \ast denote the 2-D discrete correlation and convolution operators, respectively, and where $\mathbf{e}'(i,j) = \mathbf{e}(-i,-j)$. The noise realisation \mathbf{e} is padded with at least $n_r - 1$ samples in the vertical direction and $n_c - 1$ samples in the horizontal direction by assuming periodic boundary conditions, such that \star and \ast in (8) denote 2-D circular correlation and convolution, respectively. This allows to consider only lags

$$(l,m) \in \bar{\Theta} := \{0, \dots, n_r - 1\} \times \{0, \dots, n_c - 1\}.$$

If the corruption \mathbf{e} in (1) is the realisation of a white Gaussian noise process - as in our case - it is well known that as $n \rightarrow +\infty$, the sample auto-correlation $a_{l,m}(\mathbf{e})$ vanishes for all $(l,m) \neq (0,0)$, while $a_{0,0}(\mathbf{e}) = \sigma^2$ - see, e.g., [5].

The DP exploits only the information at lag $(0,0)$. In fact, the standard deviation recovered by the residual image is required to be equal to σ . Imposing whiteness of the restoration residual by constraining the residual auto-correlation at non-zero lags to be small is a much stronger requirement.

In [6], the authors introduce the following non-negative scalar measure of whiteness $\mathcal{W} : \mathbb{R}^{n_r \times n_c} \rightarrow \mathbb{R}^+$ of noise realisation \mathbf{e} :

$$\mathcal{W}(\mathbf{e}) := \|\mathbf{e} \star \mathbf{e}\|_2^2 / \|\mathbf{e}\|_2^4 = \widetilde{\mathcal{W}}(\tilde{\mathbf{e}}), \quad (9)$$

where $\|\cdot\|_2$ denotes the Frobenius norm, while the second equality comes from Proposition 1 below, with $\widetilde{\mathcal{W}} : \mathbb{C}^{n_r \times n_c} \rightarrow \mathbb{R}^+$ the function defined in (10). Notice that the presence of the denominator in the function in (9) makes the whiteness principle completely independent of the noise level.

Proposition 1. *Let $\mathbf{e} \in \mathbb{R}^{n_r \times n_c}$ and $\tilde{\mathbf{e}} \in \mathbb{C}^{n_r \times n_c}$. Then, under the assumption of periodic boundary conditions for \mathbf{e} , the function \mathcal{W} defined in (9) satisfies:*

$$\mathcal{W}(\mathbf{e}) = \widetilde{\mathcal{W}}(\tilde{\mathbf{e}}) := \sum_{(l,m) \in \bar{\Theta}} |\tilde{e}_{l,m}|^4 / \left(\sum_{(l,m) \in \bar{\Theta}} |\tilde{e}_{l,m}|^2 \right)^2. \quad (10)$$

3 RWP for super-resolution

By now looking at (2), we observe that the nearer the super-resolved image $\mathbf{x}^*(\mu)$ is to the original image \mathbf{x} , the closer the associated residual image $\mathbf{r}^*(\mu) = \mathbf{S}\mathbf{K}\mathbf{x}^*(\mu) - \mathbf{b}$ is to the white noise realisation \mathbf{e} in (1) and, hence, the whiter is the residual image according to the scalar measure in (9).

This motivates the choice of the RWP for automatically selecting the regularisation parameter μ in variational models of the form (2), which reads:

$$\text{Select } \mu = \mu^* \text{ s.t. } \mu^* \in \arg \min_{\mu \in \mathbb{R}_*^+} W(\mu) := \mathcal{W}(\mathbf{r}^*(\mu)), \quad (11)$$

where the scalar non-negative cost function $W : \mathbb{R}_*^+ \rightarrow \mathbb{R}^+$ in (11), from now on referred to as the *residual whiteness function*, takes the following form:

$$W(\mu) = \|\mathbf{r}^*(\mu) \star \mathbf{r}^*(\mu)\|_2^2 / \|\mathbf{r}^*(\mu)\|_2^4. \quad (12)$$

Let us now give a closer look to the function in (12). First, we observe

$$\mathbf{r}^*(\mu) = \mathbf{S}\mathbf{K}\mathbf{x}^*(\mu) - \mathbf{b} = \mathbf{S}\mathbf{K}\mathbf{x}^*(\mu) - \mathbf{S}\mathbf{S}^H\mathbf{b} = \mathbf{S}\mathbf{r}_H^*(\mu),$$

where $\mathbf{r}_H^*(\mu) = \mathbf{K}\mathbf{x}^*(\mu) - \mathbf{b}_H$ is the high-resolution residual, while $\mathbf{b}_H = \mathbf{S}^H\mathbf{b}$. The denominator in (12) can be thus expressed as follows

$$\|\mathbf{r}^*(\mu)\|_2^4 = \|\mathbf{S}\mathbf{r}_H^*(\mu)\|_2^4 = \|\mathbf{S}^H\mathbf{S}\mathbf{r}_H^*(\mu)\|_2^4 = \|\mathbf{F}^H(\mathbf{F}\mathbf{S}^H\mathbf{S}\mathbf{F}^H)\mathbf{r}_H^*(\mu)\|_2^4, \quad (13)$$

where the second equality comes from recalling that \mathbf{S}^H interpolates $\mathbf{S}\mathbf{r}_H^*(\mu)$ with zeros giving null contribution when computing the norm. From Lemma 1 and by applying the Parseval's theorem, we get the following chain of equalities:

$$\begin{aligned} \|\mathbf{r}^*(\mu)\|_2^4 &= \|(1/d)\mathbf{F}^H(\mathbf{J}_{d_r} \otimes \mathbf{I}_{n_r}) \otimes (\mathbf{J}_{d_c} \otimes \mathbf{I}_{n_c})\tilde{\mathbf{r}}_H^*(\mu)\|_2^4 \\ &= \|(1/d)(\mathbf{J}_{d_r} \otimes \mathbf{I}_{n_r}) \otimes (\mathbf{J}_{d_c} \otimes \mathbf{I}_{n_c})\tilde{\mathbf{r}}_H^*(\mu)\|_2^4. \end{aligned} \quad (14)$$

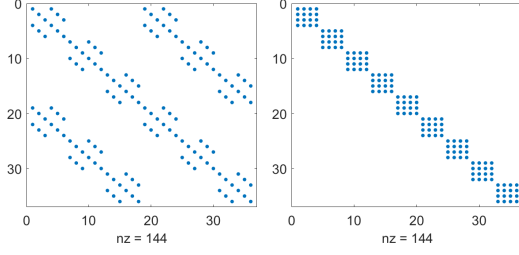


Figure 1: Structure of the matrix in (4) (left) and of the permutation induced by \mathbf{P} (right) for $n_r=n_c=3$, $d_r=d_c=2$.

The non-zero entries of the matrix introduced in Lemma 1, which are all equal to 1, are arranged along replicated patterns; this particular structure can be exploited by considering a permutation matrix $\mathbf{P} \in \mathbb{R}^{N \times N}$ such that:

$$\mathbf{P} [(\mathbf{J}_{d_r} \otimes \mathbf{I}_{n_r}) \otimes (\mathbf{J}_{d_c} \otimes \mathbf{I}_{n_c})] \mathbf{P}^T = (\mathbf{I}_n \otimes \mathbf{J}_d). \quad (15)$$

The designed permutation acts on the matrix of interest by gathering together the replicated rows and columns. In Fig. 1, we show the structure of the matrix in (4) and of the permuted matrix in (15) for $n_r=n_c=3$ and $d_r=d_c=2$.

Hence, the expression in (14) can be rewritten as

$$\|\mathbf{r}^*(\mu)\|_2^4 = \left\| \frac{1}{d} \mathbf{P} [(\mathbf{J}_{d_r} \otimes \mathbf{I}_{n_r}) \otimes (\mathbf{J}_{d_c} \otimes \mathbf{I}_{n_c})] \mathbf{P}^T \mathbf{P} \hat{\mathbf{r}}_H^*(\mu) \right\|_2^4 = \left\| \frac{1}{d} (\mathbf{I}_n \otimes \mathbf{J}_d) \hat{\mathbf{r}}_H^*(\mu) \right\|_2^4,$$

where

$$((\mathbf{I}_n \otimes \mathbf{J}_d) \hat{\mathbf{r}}_H^*(\mu))_i = \sum_{j=0}^{d-1} (\hat{\mathbf{r}}_H^*(\mu))_{\iota+j}, \quad \text{with } \iota := 1 + \left\lfloor \frac{i-1}{d} \right\rfloor d,$$

for every $i = 1, \dots, N$. The denominator in (12) can be thus expressed as

$$\|\mathbf{r}^*(\mu)\|_2^4 = \frac{1}{d^4} \left(\sum_{i=1}^N \left| \sum_{j=0}^{d-1} (\hat{\mathbf{r}}_H^*(\mu))_{\iota+j} \right|^2 \right)^2.$$

Let us now consider the numerator of the function $W(\mu)$ in (12), which, based on the definitions of auto-correlation given in (8) and of \mathbf{S}^H , reads

$$\|\mathbf{r}^*(\mu) \star \mathbf{r}^*(\mu)\|_2^2 = \|\mathbf{S} \mathbf{r}_H^*(\mu) \star \mathbf{S} \mathbf{r}_H^*(\mu)\|_2^2 = \|\mathbf{S}^H \mathbf{S} \mathbf{r}_H^*(\mu) \star \mathbf{S}^H \mathbf{S} \mathbf{r}_H^*(\mu)\|_2^2.$$

By applying again the Parseval's theorem and the convolution theorem, we get

$$\begin{aligned} \|\mathbf{r}^*(\mu) \star \mathbf{r}^*(\mu)\|_2^2 &= \|\mathbf{F} ((\mathbf{S}^H \mathbf{S} \mathbf{r}_H^*(\mu)) \star (\mathbf{S}^H \mathbf{S} \mathbf{r}_H^*(\mu)))\|_2^2 \\ &= \|\mathbf{F}(\mathbf{S}^H \mathbf{S} \mathbf{r}_H^*(\mu)) \odot \overline{\mathbf{F}(\mathbf{S}^H \mathbf{S} \mathbf{r}_H^*(\mu))}\|_2^2 \\ &= \|\mathbf{F}(\mathbf{S}^H \mathbf{S}) \mathbf{F}^H \mathbf{F} \mathbf{r}_H^*(\mu) \odot \overline{\mathbf{F}(\mathbf{S}^H \mathbf{S}) \mathbf{F}^H \mathbf{F} \mathbf{r}_H^*(\mu)}\|_2^2, \end{aligned} \quad (16)$$

where \odot denotes the Hadamard matrix product operator. The expression in (16) is manipulated by applying Lemma 1 and the permutation in (15), so as to give

$$\|\mathbf{r}^*(\mu) \star \mathbf{r}^*(\mu)\|_2^2 = \frac{1}{d^4} \sum_{i=1}^N \left| \sum_{j=0}^{d-1} (\hat{\mathbf{r}}_H^*(\mu))_{\iota+j} \right|^4. \quad (17)$$

Finally, plugging (17) and (13) into (12), we get the following form for the whiteness measure $W(\mu)$ for a super-resolution problem

$$W(\mu) = \left(\sum_{i=1}^N |w_i(\mu)|^4 \right) / \left(\sum_{i=1}^N |w_i(\mu)|^2 \right)^2, \quad w_i(\mu) = \sum_{j=0}^{d-1} (\hat{\mathbf{r}}_H(\mu))_{\iota+j}. \quad (18)$$

3.1 RWP for ℓ_2 - ℓ_2 problems in the form (2)

Here, we derive the analytical expression of the whiteness function $W(\mu)$ defined in (18) when addressing Tikhonov-regularised least squares problems as the one in (2). We start following [12] to deduce an explicit and easily-computable expression of $\mathbf{x}^*(\mu)$. By optimality, we get:

$$\mathbf{x}^*(\mu) = (\mu(\mathbf{S}\mathbf{K})^H(\mathbf{S}\mathbf{K}) + \mathbf{L}^H\mathbf{L})^{-1}(\mu(\mathbf{S}\mathbf{K})^H\mathbf{b} + \mathbf{L}^H\mathbf{v}),$$

which can be manipulated in terms of \mathbf{F} and \mathbf{F}^H to deduce

$$\begin{aligned} \mathbf{x}^*(\mu) &= (\mu\mathbf{F}^H\mathbf{F}\mathbf{K}^H\mathbf{F}^H\mathbf{F}\mathbf{S}^H\mathbf{S}\mathbf{F}^H\mathbf{F}\mathbf{K}\mathbf{F}^H\mathbf{F} + \mathbf{F}^H\mathbf{F}\mathbf{L}^T\mathbf{F}^H\mathbf{F}\mathbf{L}\mathbf{F}^H\mathbf{F})^{-1}(\mu\mathbf{K}^H\mathbf{S}^H\mathbf{b} + \mathbf{L}^H\mathbf{v}) \\ &= (\mu\mathbf{F}^H\mathbf{\Lambda}^H(\mathbf{F}\mathbf{S}^H\mathbf{S}\mathbf{F}^H)\mathbf{\Lambda}\mathbf{F} + \mathbf{F}^H\mathbf{\Gamma}^H\mathbf{\Gamma}\mathbf{F})^{-1}(\mu\mathbf{K}^H\mathbf{S}^H\mathbf{b} + \mathbf{L}^H\mathbf{v}), \end{aligned} \quad (19)$$

where $\mathbf{\Lambda}$, $\mathbf{\Gamma}$ are defined in (7). Lemma 1 provides a useful expression for the product $(\mathbf{F}\mathbf{S}^H\mathbf{S}\mathbf{F}^H)$, by which (19) becomes:

$$\begin{aligned} \mathbf{x}^*(\mu) &= \left(\frac{\mu}{d}\mathbf{F}^H\mathbf{\Lambda}^H\mathbf{P}^T(\mathbf{I}_n \otimes \mathbf{J}_d)\mathbf{P}\mathbf{\Lambda}\mathbf{F} + \mathbf{F}^H\mathbf{\Gamma}^H\mathbf{\Gamma}\mathbf{F} \right)^{-1} (\mu\mathbf{K}^H\mathbf{S}^H\mathbf{b} + \mathbf{L}^H\mathbf{v}) \\ &= \mathbf{F}^H \left(\frac{\mu}{d}\mathbf{\Lambda}^H\mathbf{P}^T(\mathbf{I}_n \otimes \mathbf{J}_d)\mathbf{P}\mathbf{\Lambda} + \mathbf{\Gamma}^H\mathbf{\Gamma} \right)^{-1} \mathbf{F}(\mu\mathbf{K}^H\mathbf{F}^H\mathbf{F}\mathbf{S}^H\mathbf{b} + \mathbf{L}^H\mathbf{F}^H\mathbf{F}\mathbf{v}) \\ &= \mathbf{F}^H \left(\frac{\mu}{d}\mathbf{\Lambda}^H\mathbf{P}^T(\mathbf{I}_n \otimes \mathbf{J}_d)\mathbf{P}\mathbf{\Lambda} + \mathbf{\Gamma}^H\mathbf{\Gamma} \right)^{-1} (\mu\mathbf{\Lambda}^H\tilde{\mathbf{b}}_H + \mathbf{\Gamma}^H\tilde{\mathbf{v}}), \end{aligned} \quad (20)$$

where $\tilde{\mathbf{b}}_H = \mathbf{F}\mathbf{b}_H = \mathbf{F}\mathbf{S}^H\mathbf{b}$ contains d replication of $\tilde{\mathbf{b}}$ - see, e.g., [8]. We now introduce the following operators

$$\underline{\mathbf{\Lambda}} := (\mathbf{I}_n \otimes \mathbf{1}_d^T)\mathbf{P}\mathbf{\Lambda} \quad \underline{\mathbf{\Lambda}}^H := \mathbf{\Lambda}^H\mathbf{P}^T(\mathbf{I}_n \otimes \mathbf{1}_d) \quad (21)$$

where $\mathbf{1}_d \in \mathbb{R}^d$ is a vector of ones. In compact form, equation (20) reads:

$$\mathbf{x}^*(\mu) = \mathbf{F}^H \left(\frac{\mu}{d}\underline{\mathbf{\Lambda}}^H\underline{\mathbf{\Lambda}} + \mathbf{\Gamma}^H\mathbf{\Gamma} \right)^{-1} (\mu\mathbf{\Lambda}^H\tilde{\mathbf{b}}_H + \mathbf{\Gamma}^H\tilde{\mathbf{v}}). \quad (22)$$

Proceeding as in [12], we can now apply the Woodbury formula (6) and perform few manipulations, so as to obtain that the expression in (22) becomes:

$$\mathbf{x}^*(\mu) = \mathbf{F}^H \left[\underline{\Psi} - \mu \underline{\Psi} \underline{\Lambda}^H \left(d\mathbf{I} + \mu \underline{\Lambda} \underline{\Psi} \underline{\Lambda}^H \right)^{-1} \underline{\Lambda} \underline{\Psi} \right] (\mu \underline{\Lambda}^H \tilde{\mathbf{b}}_H + \Gamma^H \tilde{\mathbf{v}}), \quad (23)$$

whence the Fourier transform of the high resolution residual $\mathbf{r}_H^*(\mu) = \mathbf{K} \mathbf{x}^*(\mu) - \mathbf{b}$, with $\mathbf{x}^*(\mu)$ given in (23), can be written as

$$\tilde{\mathbf{r}}_H^*(\mu) = \underline{\Lambda} \left[\underline{\Psi} - \mu \underline{\Psi} \underline{\Lambda}^H \left(d\mathbf{I} + \mu \underline{\Lambda} \underline{\Psi} \underline{\Lambda}^H \right)^{-1} \underline{\Lambda} \underline{\Psi} \right] (\mu \underline{\Lambda}^H \tilde{\mathbf{b}}_H + \Gamma^H \tilde{\mathbf{v}}) - \tilde{\mathbf{b}}_H, \quad (24)$$

where $\underline{\Psi} = (\Gamma^H \Gamma + \epsilon)^{-1}$ and the parameter $0 < \epsilon \ll 1$ guarantees the inversion of $\Gamma^H \Gamma$. Recalling Lemma 2 and the property (15), we prove the following result.

Proposition 2. *Let $\Phi \in \mathbb{R}^{n \times n}$ be a diagonal matrix and consider the matrix $\underline{\Lambda}$ defined in (21). Then, the following equality holds:*

$$\underline{\Lambda}^H \Phi \underline{\Lambda} = \mathbf{P}^T (\Phi \otimes \mathbf{I}_d) \mathbf{P} \underline{\Lambda}^H \underline{\Lambda}.$$

Proof. Recalling property (5) in Lemma 2, we get the following chain of equalities

$$\begin{aligned} \underline{\Lambda}^H \Phi \underline{\Lambda} &= \underline{\Lambda}^H \mathbf{P}^T (\mathbf{I}_n \otimes \mathbf{1}_d) \Phi (\mathbf{I}_n \otimes \mathbf{1}_d^T) \mathbf{P} \underline{\Lambda} = \underline{\Lambda}^H \mathbf{P}^T (\mathbf{I}_n \otimes \mathbf{1}_d) (\Phi \otimes \mathbf{1}_d^T) \mathbf{P} \underline{\Lambda} \\ &= \underline{\Lambda}^H \mathbf{P}^T (\mathbf{I}_n \Phi \otimes \mathbf{1}_d \mathbf{1}_d^T) \mathbf{P} \underline{\Lambda} = \underline{\Lambda}^H \mathbf{P}^T (\Phi \mathbf{I}_n \otimes \mathbf{J}_d) \mathbf{P} \underline{\Lambda} \\ &= \underline{\Lambda}^H \mathbf{P}^T (\Phi \mathbf{I}_n \otimes \mathbf{I}_d \mathbf{J}_d) \mathbf{P} \underline{\Lambda} = \underline{\Lambda}^H \mathbf{P}^T (\Phi \otimes \mathbf{I}_d) (\mathbf{I}_n \otimes \mathbf{J}_d) \mathbf{P} \underline{\Lambda} \\ &= \underline{\Lambda}^H \mathbf{P}^T (\Phi \otimes \mathbf{I}_d) \mathbf{P} \mathbf{P}^T (\mathbf{I}_n \otimes \mathbf{J}_d) \mathbf{P} \underline{\Lambda}, \end{aligned}$$

where the sparse block-diagonal matrix $\mathbf{P}^T (\Phi \otimes \mathbf{I}_d) \mathbf{P} \in \mathbb{R}^{N \times N}$ commutes with $\underline{\Lambda}^H$, so that $\underline{\Lambda}^H \mathbf{P}^T (\Phi \otimes \mathbf{I}_d) \mathbf{P} = \mathbf{P}^T (\Phi \otimes \mathbf{I}_d) \mathbf{P} \underline{\Lambda}^H$. Recalling (21), this yields:

$$\underline{\Lambda}^H \Phi \underline{\Lambda} = \mathbf{P}^T (\Phi \otimes \mathbf{I}_d) \mathbf{P} \underline{\Lambda}^H \underline{\Lambda},$$

which completes the proof. \square \square

Corollary 1. *Let $\Phi = (d\mathbf{I} + \mu \underline{\Lambda} \underline{\Psi} \underline{\Lambda}^H)^{-1}$. Then, the expression in (24) turns into*

$$\tilde{\mathbf{r}}_H^*(\mu) = \underline{\Lambda} \left[\underline{\Psi} - \mu \underline{\Psi} \mathbf{P}^T \left((d\mathbf{I} + \mu \underline{\Lambda} \underline{\Psi} \underline{\Lambda}^H)^{-1} \otimes \mathbf{I}_d \right) \mathbf{P} \underline{\Lambda}^H \underline{\Lambda} \underline{\Psi} \right] (\mu \underline{\Lambda}^H \tilde{\mathbf{b}}_H + \Gamma^H \tilde{\mathbf{v}}) - \tilde{\mathbf{b}}_H.$$

Proof. We first notice that

$$\underline{\Lambda} \underline{\Psi} \underline{\Lambda}^H = (\mathbf{I}_n \otimes \mathbf{1}_d^T) \widehat{\underline{\Lambda} \underline{\Psi} \underline{\Lambda}^H} (\mathbf{I}_n \otimes \mathbf{1}_d), \quad (25)$$

is diagonal as $\widehat{\underline{\Lambda} \underline{\Psi} \underline{\Lambda}^H} = \mathbf{P} \underline{\Lambda} \underline{\Psi} \underline{\Lambda}^H \mathbf{P}^T$ is. The matrix in (25) can thus be written as

$$\underline{\Lambda} \underline{\Psi} \underline{\Lambda}^H = \text{diag}(\omega_1, \dots, \omega_n), \quad \omega_i = \sum_{j=0}^{d-1} \frac{|\hat{\lambda}_{i+j}|^2}{|\hat{\gamma}_{i+j}|^2 + \epsilon}.$$

Hence, since Φ is the inverse of the sum of two diagonal matrices, it is diagonal so we can apply Proposition 2 and deduce the thesis. \square \square

Recalling now the action of the permutation matrix \mathbf{P} on vectors, we have that the product $\hat{\mathbf{r}}_H^*(\mu) = \mathbf{P}\tilde{\mathbf{r}}_H^*(\mu)$ reads

$$\hat{\mathbf{r}}_H^*(\mu) = \left[\widehat{\underline{\Lambda}\Psi} - \mu \widehat{\underline{\Lambda}\Psi} \left((d\mathbf{I} + \mu \underline{\Lambda}\Psi \underline{\Lambda}^H)^{-1} \otimes \mathbf{I}_d \right) \widehat{\underline{\Lambda}^H \underline{\Lambda}\Psi} \right] (\mu \check{\underline{\Lambda}} \hat{\mathbf{b}}_H + \check{\Gamma} \hat{\mathbf{v}}) - \hat{\mathbf{b}}_H, \quad (26)$$

where the matrix $\widehat{\underline{\Lambda}^H \underline{\Lambda}\Psi} = \mathbf{P} \underline{\Lambda}^H \underline{\Lambda} \Psi \mathbf{P}^T (\mathbf{I}_n \otimes \mathbf{J}_d) \mathbf{P} \underline{\Lambda} \Psi \mathbf{P}^T$ acts on $\mathbf{g} \in \mathbb{R}^N$ as

$$(\widehat{\underline{\Lambda}^H \underline{\Lambda}\Psi} \mathbf{g})_i = \tilde{\lambda}_i \sum_{j=0}^{d-1} \frac{\hat{\lambda}_{\iota+j}}{|\hat{\gamma}_{\iota+j}|^2 + \epsilon} g_{\iota+j}.$$

Combining altogether, we finally deduce:

$$\begin{aligned} \hat{\mathbf{r}}_H^*(\mu) &= \mu \widehat{\underline{\Lambda}\Psi} \check{\underline{\Lambda}} \hat{\mathbf{b}}_H + \widehat{\underline{\Lambda}\Psi} \check{\Gamma} \hat{\mathbf{v}} - \mu^2 \widehat{\underline{\Lambda}\Psi} \left[(d\mathbf{I} + \mu \underline{\Lambda}\Psi \underline{\Lambda}^H)^{-1} \otimes \mathbf{I}_d \right] \widehat{\underline{\Lambda}^H \underline{\Lambda}\Psi} \check{\underline{\Lambda}} \hat{\mathbf{b}}_H \\ &\quad - \mu \widehat{\underline{\Lambda}\Psi} \left[(d\mathbf{I} + \mu \underline{\Lambda}\Psi \underline{\Lambda}^H)^{-1} \otimes \mathbf{I}_d \right] \widehat{\underline{\Lambda}^H \underline{\Lambda}\Psi} \check{\Gamma} \hat{\mathbf{v}} - \hat{\mathbf{b}}_H, \end{aligned}$$

whence we can explicitly compute the expression for each component $i = 1, \dots, n$:

$$\begin{aligned} (\hat{\mathbf{r}}_H^*(\mu))_i &= \mu \left[\frac{|\hat{\lambda}_i|^2}{|\hat{\gamma}_i|^2 + \epsilon} \hat{b}_{H,i} \right] + \frac{\hat{\lambda}_i \bar{\gamma}_i \hat{v}_i}{|\hat{\gamma}_i|^2 + \epsilon} - \left[\mu^2 \sum_{j=0}^{d-1} \frac{|\hat{\lambda}_{\iota+j}|^2 \hat{b}_{H,\iota+j}}{|\hat{\gamma}_{\iota+j}|^2 + \epsilon} \right. \\ &\quad \left. + \mu \sum_{j=0}^{d-1} \frac{\hat{\lambda}_{\iota+j} \bar{\gamma}_{\iota+j} \hat{v}_{\iota+j}}{|\hat{\gamma}_{\iota+j}|^2 + \epsilon} \right] \frac{|\hat{\lambda}_i|^2}{|\hat{\gamma}_i|^2 + \epsilon} \left(d + \mu \sum_{j=0}^{d-1} \frac{|\hat{\lambda}_{\iota+j}|^2}{|\hat{\gamma}_{\iota+j}|^2 + \epsilon} \right)^{-1} - \hat{b}_{H,i}. \end{aligned}$$

We can thus deduce the following expression of the terms in formula (18)

$$\begin{aligned} \sum_{j=0}^{d-1} (\hat{\mathbf{r}}_H^*(\mu))_{\iota+j} &= \frac{1}{d + \mu \sum_{j=0}^{d-1} \frac{|\hat{\lambda}_{\iota+j}|^2}{|\hat{\gamma}_{\iota+j}|^2 + \epsilon}} \left[\mu \left(d \sum_{j=0}^{d-1} \frac{|\hat{\lambda}_{\iota+j}|^2}{|\hat{\gamma}_{\iota+j}|^2 + \epsilon} \hat{b}_{H,\iota+j} \right. \right. \\ &\quad \left. \left. - \sum_{j=0}^{d-1} \hat{b}_{H,\iota+j} \sum_{j=0}^{d-1} \frac{|\hat{\lambda}_{\iota+j}|^2}{|\hat{\gamma}_{\iota+j}|^2 + \epsilon} \right) + d \left(\sum_{j=0}^{d-1} \frac{\hat{\lambda}_{\iota+j} \bar{\gamma}_{\iota+j} \hat{v}_{\iota+j}}{|\hat{\gamma}_{\iota+j}|^2 + \epsilon} - \sum_{j=0}^{d-1} \hat{b}_{H,\iota+j} \right) \right]. \end{aligned} \quad (27)$$

In light of its replicating structure, we observe that the action of the permutation \mathbf{P} on $\tilde{\mathbf{b}}_H$ will cluster the identical entries, so that the $\hat{b}_{H,\iota+j}$ can be written as the mean of the set of d values $\{\hat{b}_{H,\iota}, \dots, \hat{b}_{H,\iota+d-1}\}$. This allows to simplify formula (27) as the difference in the first bracket vanishes. By now setting

$$\eta_i := \frac{1}{d} \sum_{j=0}^{d-1} \frac{|\hat{\lambda}_{\iota+j}|^2}{|\hat{\gamma}_{\iota+j}|^2 + \epsilon}, \quad \varrho_i := \sum_{j=0}^{d-1} \hat{b}_{H,\iota+j}, \quad \nu_i := \sum_{j=0}^{d-1} \frac{\hat{\lambda}_{\iota+j} \bar{\gamma}_{\iota+j} \hat{v}_{\iota+j}}{|\hat{\gamma}_{\iota+j}|^2 + \epsilon},$$

which can all be computed beforehand. Plugging (27) into (18) we finally get

$$W(\mu) = \left(\sum_{i=1}^N \left| \frac{\nu_i - \varrho_i}{1 + \eta_i \mu} \right|^4 \right) / \left(\sum_{i=1}^N \left| \frac{\nu_i - \varrho_i}{1 + \eta_i \mu} \right|^2 \right)^2. \quad (28)$$

Note that when $d = 1$, i.e. when no decimation is considered, this formula corresponds exactly to the one considered in [6] in the context of image deblurring.

According to the RWP, the optimal μ^* is selected as the one minimising the whiteness measure function in (28). We remark that the action of the permutation matrix \mathbf{P} can be efficiently replicated without deriving its explicit expression; as a result, the overall computational cost for the evaluation of $W(\mu)$ amounts to $O(N \log N)$, namely the cost of the 2D fast Fourier transform and of its inverse, and the value μ^* can be efficiently detected via grid-search. Finally, the optimal μ^* is used for the computation of the reconstruction $\mathbf{x}^*(\mu^*)$ based on (23).

The main steps of the proposed procedure are summarised in Algorithm 1.

Algorithm 1: SR for (2) with automatic parameter selection via RWP

inputs: observed image $\mathbf{b} \in \mathbb{R}^n$, forward model operator $\mathbf{K} \in \mathbb{R}^{nd \times nd}$,
down-sampling operator $\mathbf{S} \in \mathbb{R}^{n \times nd}$

- **Compute Fourier diagonalisations:** $\mathbf{\Lambda} = \mathbf{F}\mathbf{K}\mathbf{F}^H$, $\mathbf{\Gamma} = \mathbf{F}\mathbf{L}\mathbf{F}^H$
 - **Compute matrices:** $\mathbf{\underline{\Lambda}} = (\mathbf{I}_n \otimes \mathbf{1}_d^T)\mathbf{P}\mathbf{\Lambda}$, $\mathbf{\Psi} = (\mathbf{\Gamma}^H\mathbf{\Gamma} + \epsilon)^{-1}$
 - **Residual whiteness principle for the selection of μ^* :**
 - Compute $W(\mu)$ in (28) for different values of μ , based on Corollary 1 and (26)
 - Select $\mu^* \in \arg \min W(\mu)$
 - **Compute the reconstruction:** $\mathbf{x}^*(\mu^*)$ by (23)
-

4 Numerical results

We evaluate the proposed RWP-based automatic procedure for selecting the regularisation parameter μ in variational models of the form (2) when $\mathbf{v} = \mathbf{0}_N$ and $\mathbf{L} = \mathbf{D} := (\mathbf{D}_h^T, \mathbf{D}_v^T)^T \in \mathbb{R}^{2N \times N}$, with $\mathbf{D}_h, \mathbf{D}_v \in \mathbb{R}^{N \times N}$ representing the finite difference operators discretising the first-order horizontal and vertical partial derivatives, respectively. Note that \mathbf{D} verifies assumptions (A3)-(A4) in Sec. 1.1.

Our goal is to highlight that the proposed RWP selects a regularisation parameter value μ^* yielding high quality restorations. The RWP is compared with the DP, defined in (3) when $\tau = 1$. There is a one-to-one relationship between the μ -value and the norm of the associated residual image. Hence, in all the presented results we will substitute the μ -values with the corresponding τ -values, with τ defined according to (3) by $\tau^*(\mu) := \|\mathbf{S}\mathbf{H}\mathbf{x}^*(\mu) - \mathbf{b}\|_2 / (\sqrt{n}\sigma)$.

The quality of the restorations \mathbf{x}^* , for different values of τ^* , with respect to the original undecimated image \mathbf{x} , will be assessed by means of three scalar

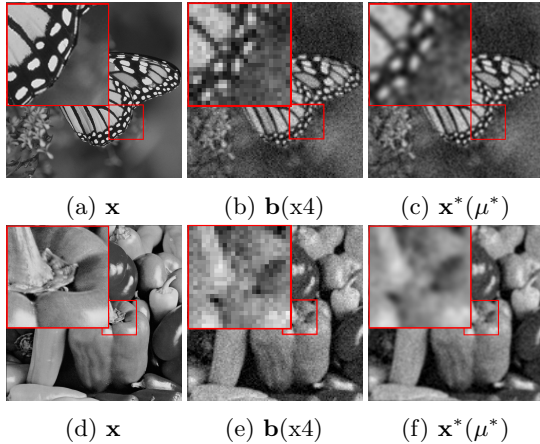


Figure 2: From left to right: original \mathbf{x} , observed \mathbf{b} , our reconstruction $\mathbf{x}^*(\mu^*)$ for monarch (top) and peppers (bottom).

| monarch | | | | |
|--------------------|---------|--------|--------|--------|
| | PSNR | ISNR | SSIM | |
| \mathbf{x}^* | 21.4050 | 1.3452 | 0.6736 | Test 1 |
| $\bar{\mathbf{b}}$ | 20.0598 | - | 0.5435 | |
| \mathbf{x}^* | 18.9277 | 1.1561 | 0.6297 | Test 2 |
| $\bar{\mathbf{b}}$ | 17.7716 | - | 0.3105 | |

| peppers | | | | |
|--------------------|---------|--------|--------|--------|
| | PSNR | ISNR | SSIM | |
| \mathbf{x}^* | 23.5674 | 1.9147 | 0.6757 | Test 1 |
| $\bar{\mathbf{b}}$ | 21.6526 | - | 0.5187 | |
| \mathbf{x}^* | 21.3034 | 2.6078 | 0.6240 | Test 2 |
| $\bar{\mathbf{b}}$ | 18.6956 | - | 0.3032 | |

Table 1: Achieved PSNR, ISNR, SSIM values for monarch and peppers for the two degradation settings.

measures, namely the Structural Similarity Index (SSIM) [11], the Peak-Signal-to-Noise-Ratio (PSNR) and the Improved-Signal-to-Noise Ratio (ISNR), defined by $\text{PSNR} = 20 \log_{10}(\sqrt{N} \max(\mathbf{x}, \mathbf{x}^*) / \|\mathbf{x} - \mathbf{x}^*\|_2)$ and $\text{ISNR} = 10 \log_{10}(\|\mathbf{x} - \bar{\mathbf{b}}\|_2 / \|\mathbf{x} - \mathbf{x}^*\|_2)$, respectively, with $\max(\mathbf{x}, \mathbf{x}^*)$ representing the largest value of \mathbf{x} and \mathbf{x}^* , while $\bar{\mathbf{b}}$ denotes the bicubic interpolation of \mathbf{b} .

We consider two test images of size 512×512 with pixel values normalised in $[0, 1]$, namely **monarch** and **peppers**, shown in Figs. 2a-2d, respectively. The decimation factors along the rows and the columns of the original images are set as $d_c = d_r = 4$. As a first example, the original test images are corrupted by Gaussian blur, generated by the Matlab routine `fspecial` with input parameters `band=9` and `sigma=2`. The `band` parameter represents the side length (in pixels) of the square support of the kernel, whereas `sigma` is the standard deviation (in pixels) of the isotropic bivariate Gaussian distribution defining the kernel in the continuous setting. Finally, the decimated and blurred images are corrupted by an AWGN with standard deviation $\sigma = 0.05$. The observed data for the test images **monarch** and **peppers** are displayed in Fig. 2b-2e, respectively.

In Figs. 3a, we report the behavior of the whiteness measure $W(\mu)$ as a function of $\tau^*(\mu)$ for the test images **monarch** (solid blue line) and **peppers** (solid black line), respectively. The plotted values have been obtained by solving the model (2) for a fine grid of different μ values, and then computing for each μ the associated $\tau^*(\mu)$ and $W(\mu)$. The optimal τ^* s corresponding to μ^* s are indicated by the vertical dashed magenta and green lines for **monarch** ($\tau^*(\mu^*) = 0.9398$) and **peppers** ($\tau^*(\mu^*) = 0.9633$), respectively, while $\tau = 1$, representing the DP,

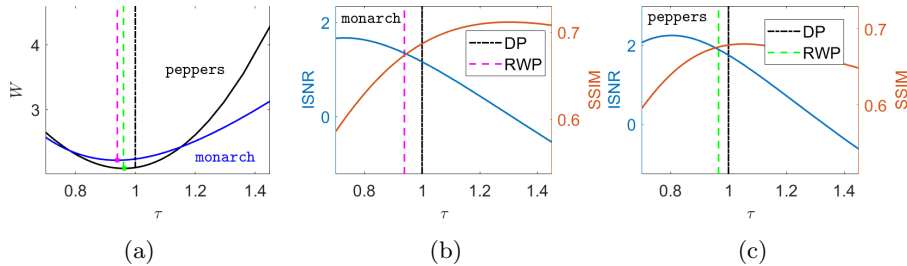


Figure 3: Whiteness measure functions (left column) and ISNR/SSIM values for different τ s (second column) for the `monarch` and `peppers` test images.

is depicted by the vertical dotted black line. Notice that the whiteness curves computed *a posteriori* admit a minimiser over the considered domain which coincides with the τ^* selected by the RWP. Moreover, the proximity of the optimal τ^* s to 1 indicates that the noise level estimated starting from $\mathbf{r}^*(\mu)$ is close to the true one.

In Figs. 3b-3c, we graphic the achieved ISNR and SSIM for the two test images. Note that the RWP tends to automatically select a μ -value returning the best trade-off between the two quality measures. The reconstructed $\mathbf{x}^*(\mu^*)$ for the two test images are shown in Figs. 2c-2f. Finally, the PSNR, ISNR and SSIM values achieved by the proposed strategy are reported in the top part of Tab. 1 (Test 1), together with the PSNR and SSIM of the bicubic interpolation.

As a second example, we perform the same reconstructions with different degradation levels. More specifically, we consider a Gaussian blur with parameters `band` = 13, `sigma` = 3, and AWGN with standard deviation $\sigma = 0.1$. The achieved quality measures are reported in the bottom part of Tab.1 (Test 2). In this case, the RWP automatically selects an optimal τ^* corresponding to a very accurate estimate of the original noise standard deviation, namely $\tau^*(\mu^*) = 0.9938$ for `monarch` and $\tau^*(\mu^*) = 0.9915$ for `peppers`.

5 Conclusions and outlook

We extended the residual whiteness principle introduced in [6] for automatic parameter selection with unknown noise level in image deblurring to image super-resolution problems solved by generalised Tikhonov regularisation models in the form (2) whose solution can be efficiently computed by means of the approach outlined in [12]. By exploiting carefully technical properties of the operators involved in the model in the frequency domain, a compact formula for the whiteness measure can be found. Its minimisation provides an accurate estimate of the unknown noise level. As a future work, we plan to explicitly formalise the extension of the RWP to non-smooth super-resolution models as well as to explicitly tackle the minimisation of the whiteness measure with more sophisticated techniques.

References

- [1] A. Z. Chen and B. X. Huo and C. Y. Wen (2013) Adaptive regularization for color image restoration using discrepancy principle. In ICSPCC 2013:1-6.
- [2] C. Clason (1996) Regularization of Inverse Problems. Kluwer, Dordrecht.
- [3] P. Craven and G. Wahba (1978) Smoothing noisy data with spline functions. *Numer. Math.*, 31:377-403.
- [4] P. Hansen (1987) Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion.
- [5] A. Lanza and S. Morigi and F. Sciacchitano and F. Sgallari (2018) Whiteness Constraints in a Unified Variational Framework for Image Restoration. *J. Math. Imaging Vis.*, 60:1503-1526.
- [6] A. Lanza and M. Pragliola and F. Sgallari (2020) Residual whiteness principle for parameter-free image restoration. *Electron. Trans. Numer. Anal.*, 53:329-351.
- [7] P. Riot (2018) Blancheur du résidu pour le débruitage d'image. PhD Thesis.
- [8] M. D. Robinson and S. Farsiu and J. Y. Lo and P. Milanfar and C. Toth (2007) Efficient Registration of Aliased X-Ray Images. *ACSSC*:215-219.
- [9] A. Toma and B. Sixou and F. Peyrin (2015) Iterative choice of the optimal regularization parameter in TV image restoration. *Inverse Problems & Imaging*, 9:1171.
- [10] N. K. Tuador and D. Pham and J. Michetti and A. Basarab and D. Kouamé (2020) A Novel Fast 3D Single Image Super-Resolution Algorithm. *ArXiv*, abs/2010.15491.
- [11] Z. Wang and A. Bovik and H. R. Sheikh and E. P. Simoncelli (2004) Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13:600-612.
- [12] N. Zhao and Q. Wei and A. Basarab and N. Dobigeon and D. Kouamé and J. Tourneret (2016) Fast Single Image Super-Resolution Using a New Analytical Solution for ℓ_2 - ℓ_2 Problems. *IEEE Trans. Image Process.*, 25:3683-3697.