MULTISCALE MODEL. SIMUL.
Vol. 4, No. 2, pp. 390–423

IMAGE DECOMPOSITION USING TOTAL VARIATION AND div(BMO)*

TRIET M. LE† AND LUMINITA A. VESE†

Abstract. This paper is devoted to the decomposition of an image f into $u + v$, with u a piecewise-smooth or "cartoon" component, and v an oscillatory component (texture or noise), in a variational approach. Meyer [*Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, Univ. Lecture Ser. 22, AMS, Providence, RI, 2001] proposed refinements of the total variation model (Rudin, Osher, and Fatemi [*Phys. D*, 60 (1992), pp. 259–268]) that better represent the oscillatory part v: the spaces of generalized functions $G = \text{div}(L^\infty)$ and $F = \text{div}(BMO)$ (this last space arises in the study of Navier–Stokes equations; see Koch and Tataru [*Adv. Math.*, 157 (2001), pp. 22–35]) have been proposed to model v, instead of the standard L^2 space, while keeping u a function of bounded variation. Mumford and Gidas [*Quart. Appl. Math.*, 59 (2001), pp. 85–111] also show that natural images can be seen as samples of scale-invariant probability distributions that are supported on distributions only and not on sets of functions. However, there is no simple solution to obtain in practice such decompositions $f = u + v$ when working with G or F. In earlier works [L. Vese and S. Osher, *J. Sci. Comput.*, 19 (2003), pp. 553–572], [L. A. Vese and S. J. Osher, *J. Math. Imaging Vision*, 20 (2004), pp. 7–18], [S. Osher, A. Solé, and L. Vese, *Multiscale Model. Simul.*, 1 (2003), pp. 349–370], the authors have proposed approximations to the (BV, G) decomposition model, where the L^∞ space has been substituted by L^p, $1 \le p < \infty$. In the present paper, we introduce energy minimization models to compute (BV, F) decompositions, and as a by-product we also introduce a simple model to realize the (BV, G) decomposition. In particular, we investigate several methods for the computation of the BMO norm of a function in practice. Theoretical, experimental results and comparisons to validate the proposed new methods are presented.

Key words. functional minimization, nonlinear partial differential equation, bounded variation, bounded mean oscillation, oscillatory function, image decomposition, cartoon and texture, denoising

AMS subject classifications. 35, 49, 65

DOI. 10.1137/040610052

1. Introduction and motivations.

In what follows, we assume that a given grayscale image can be represented by a function (or sometimes distribution) f, defined on an open, bounded, and connected subset Ω of $\mathbb{R}^2$, with Lipschitz boundary $\partial\Omega$. In general, Ω is a rectangle in the plane. Sometimes, we may assume that the image f is defined everywhere in the plane (obtained by extension). We limit our presentation to the two-dimensional case, but our results hold in any dimension.

We are interested in decomposing f into $u+v$ via an energy minimization problem

$$\inf_{(u,v)\in X_1\times X_2} \{\mathcal{K}(u,v) = F_1(u) + \lambda F_2(v) : f = u + v\},$$

where $F_1, F_2 \ge 0$ are functionals and X_1, X_2 are spaces of functions or distributions such that $X_1 = \{u : F_1(u) < \infty\}$, $X_2 = \{v : F_2(v) < \infty\}$. It is assumed that $f \in X_1 + X_2$. The constant $\lambda > 0$ is a tuning parameter. Usually, F_1 and F_2 are norms or seminorms of functional spaces arising in image analysis (i.e., $F_i(\cdot) = \|\cdot\|_{X_i}$).

*Received by the editors June 16, 2004; accepted for publication (in revised form) March 2, 2005; published electronically June 27, 2005. This work was supported in part by an Alfred P. Sloan fellowship, by the UCLA Institute of Pure and Applied Mathematics, by the National Science Foundation (grant ITR/ACI-0113439), and by the National Institute of Health through the NIH Roadmap for Medical Research (grant U54 RR021813).

http://www.siam.org/journals/mms/4-2/61005.html

†Department of Mathematics, University of California, Los Angeles, 405 Hilgard Ave., Los Angeles, CA 90095-1555 (tle@math.ucla.edu, lvese@math.ucla.edu).

An important problem in image analysis is to separate different features in images. For instance, in image denoising, f is the observed noisy version of the true unknown image u, while v represents additive Gaussian noise of zero mean. Often in this case, $X_1 \subset X_2$, $f \in X_2$, and X_1 is a space of functions "smoother" or less oscillating than those in X_2. However, sharp edges or boundaries have to be represented in u. Another related problem is the separation of the geometric (cartoon) component u of f from the oscillatory component v, representing texture or noise, of zero mean. In other cases, u can be seen as a geometric or structure component of f, while v is clutter; see [47]. A good model for $\mathcal{K}$ is given by a choice of X_1 and X_2 so that with the above given properties of u and v, the (semi-) norms $F_1(u) = \|u\|_{X_1}$ and $F_2(v) = \|v\|_{X_2}$ are small. We give here two examples of image decomposition models by variational methods that are most related with our framework. However, many other previous work (variational or nonvariational) can be seen as decompositions of f into $u + v$. In the Mumford and Shah model for image segmentation [30], $f \in L^\infty(\Omega) \subset L^2(\Omega)$ is split into $u \in SBV(\Omega)$ [28], [3] (a piecewise-smooth function with its discontinuity set J_u composed of a union of curves of total finite length), and $v = f - u \in L^2(\Omega)$ represents noise or texture. The problem in the weak formulation is [30], [28]

$$\inf_{(u,v)\in SBV(\Omega)\times L^2(\Omega)} \left\{ \int_{\Omega\setminus J_u} |\nabla u|^2 + \alpha \mathcal{H}^1(J_u) + \beta \|v\|^2_{L^2(\Omega)}, f = u + v \right\}, \tag{1}$$

where $\mathcal{H}^1$ denotes the one-dimensional Hausdorff measure, and $\alpha, \beta > 0$ are tuning parameters. With the above notation, the first two terms in the energy from (1) compose $F_1(u)$, while the third term makes $F_2(v)$. A related decomposition is obtained by the total variation minimization model of Rudin, Osher, and Fatemi [34] for image denoising, where $SBV(\Omega)$ is substituted by the slightly larger space $BV(\Omega)$ of functions of bounded variation that is defined by [18], [4], [5].

DEFINITION 1.1. *Let $u \in L^1(\Omega)$; we say that u is a function of bounded variation in Ω if the distributional derivative of u is representable by a finite Radon measure in Ω, i.e., if*

$$\int_\Omega u \frac{\partial \phi}{\partial x_i} dx = -\int_\Omega \phi D_i u \quad \forall \phi \in C_c^1(\Omega), \quad i = 1, 2,$$

for some $\mathbb{R}^2$-valued measure $Du = (D_1 u, D_2 u)$ in Ω. The vector space of all functions of bounded variation in Ω is denoted by $BV(\Omega)$.

Another equivalent definition of the space $BV(\Omega)$ (as a dual space) is obtained by the following definition.

DEFINITION 1.2. *Let $u \in L^1(\Omega)$. The variation of u in Ω is defined by*

$$V(u, \Omega) := \sup \left\{ \int_\Omega u(\operatorname{div} \vec{g}) dx : \ \vec{g} \in \left[C_c^1(\Omega)\right]^2, \ \|\vec{g}\|_{L^\infty(\Omega)} \le 1 \right\}.$$

PROPOSITION 1.3. *Let $u \in L^1(\Omega)$. Then $u \in BV(\Omega)$ if and only if $V(u, \Omega) < \infty$. In addition, $V(u, \Omega) = |Du|(\Omega)$ for any $u \in BV(\Omega)$.*

Note that when $u \in W^{1,1}(\Omega)$, then $Du = \nabla u dx$, but the inclusion $W^{1,1}(\Omega) \subset BV(\Omega)$ is strict. However, by slight abuse of notation, we will sometimes use $|Du|(\Omega) = \int_\Omega |\nabla u| dx = |u|_{BV(\Omega)}$ for $u \in BV(\Omega)$, where $|u|_{BV(\Omega)}$ is the seminorm. Equipped with $\|u\|_{BV(\Omega)} := |Du|(\Omega) + \|u\|_{L^1(\Omega)}$, $BV(\Omega)$ becomes a Banach space.

The Rudin–Osher–Fatemi (ROF) decomposition model can be defined as [34]

$$\inf_{(u,v)\in BV(\Omega)\times L^2(\Omega)} \big\{\mathcal{J}(u,v) = |u|_{BV(\Omega)} + \lambda\|v\|^2_{L^2(\Omega)}, \quad f = u + v\big\}, \tag{2}$$

where $\lambda > 0$ is a tuning parameter. In the original total variation model, v represents additive Gaussian noise of zero mean. This model provides a unique $(BV(\Omega), L^2(\Omega))$ decomposition of $f \in L^2(\Omega)$ for each $\lambda > 0$ (see [10] or [43] for a more general case). The model is convex, easy to solve in practice, and denoises well piecewise-constant images while preserving edges. However, it has some limitations. For instance, if f is the characteristic function of a smooth set E of finite perimeter, the model should produce $u = f$, $v = 0$. But this is not true for any finite value of λ [27], [39], [5]. Cartoon or BV pieces of f are sent to v, and the model does not always represent well texture or oscillatory details, as we will see later. In [40], the authors have proposed a hierarchical multiscale $(BV(\Omega), L^2(\Omega))$ decomposition to reduce such artifacts. Also, in [23], [2], it has been shown that natural images are not well represented by functions of bounded variation.

We recall the following definition.

DEFINITION 1.4. *Let $\mathcal{D}(\Omega)$ be the set of test functions in Ω, i.e., the set of all functions ϕ on Ω that are infinitely differentiable and, together with all their derivatives, are rapidly decreasing (i.e., remain bounded when multiplied by arbitrary polynomials) near the boundary $\partial\Omega$. The set of all distributions (linear continuous functionals on $\mathcal{D}(\Omega)$) is denoted by $\mathcal{D}'(\Omega)$.*

Here we are interested in a better choice for the oscillatory component v or for the space X_2, which has to give small norms for oscillatory functions, while keeping $X_1 = BV(\Omega)$. Our discussion follows Meyer [27], together with the motivations from Mumford and Gidas [29]. The idea is to use weaker norms for the oscillatory component v, instead of the $L^2(\Omega)$ norm, and this can be done by the use of generalized functions. For instance, Meyer suggests the use of $v \in (BV(\Omega))'$, the dual of the $BV(\Omega)$ space, having the inclusions $BV(\Omega) \subset L^2(\Omega) \subset (BV(\Omega))'$. However, there is no known integral representation of continuous linear functionals on $BV(\Omega)$. There is a result that describes the dual of the $SBV(\Omega)$ space by De Pauw [17], but it leads to a complicated representation. To overcome this, Meyer [27] suggests approximating $(BV(\Omega))'$ by another slightly larger space, the dual $(W_0^{1,1}(\Omega))' = W^{-1,\infty}(\Omega)$. This is equivalent with the following space of distributions [1], [27].

DEFINITION 1.5. *Let $G = G(\Omega)$ consist of distributions T in $\mathcal{D}'(\Omega)$ which can be written as*

$$T = \operatorname{div}(\vec{g}) \text{ in } \mathcal{D}'(\Omega), \quad \vec{g} = (g_1, g_2) \in (L^\infty(\Omega))^2;$$

i.e., $T(\phi) = -\int_\Omega (g_1 \frac{\partial\phi}{\partial x} + g_2 \frac{\partial\phi}{\partial y})$ for any $\phi \in \mathcal{D}(\Omega)$.

Define $\|.\|_G$ on G by

$$\|T\|_G = \inf\Big\{\|\sqrt{(g_1)^2 + (g_2)^2}\|_{L^\infty(\Omega)} : \ T = \operatorname{div}(\vec{g}) \text{ in } \mathcal{D}'(\Omega), \ \vec{g} \in L^\infty(\Omega, \mathbb{R}^2)\Big\}.$$

We recall that $W_0^{1,1}(\Omega)$ is the closure of $C_0^\infty(\Omega)$ in the space $W^{1,1}(\Omega)$. Functions in $W_0^{1,1}(\Omega)$ have zero trace on $\partial\Omega$. The space G is a Banach space, because it is isometrically isomorphic with the dual space (equipped with the dual norm) of the normed space $W_0^{1,1}(\Omega)$ equipped with $\|u\|_{W_0^{1,1}(\Omega)} = \int_\Omega |\nabla u|$. We denote this dual space in the usual way by $W^{-1,\infty}(\Omega)$.

REMARK 1. *When Ω is open, bounded, and connected in $\mathbb{R}^2$ with Lipschitz boundary, under the additional (technical) assumption $f \in L^2(\Omega)$, we have the following: if f is decomposed into $u+v$ by a (BV, X_2) model, with $u \in BV(\Omega) \subset L^2(\Omega)$ and $v = f - u \in X_2(\Omega)$, then we must have $v \in L^2(\Omega)$. Since v corresponds to additive noise and texture of zero mean $\int_\Omega v = 0$, Aubert and Aujol* [6]*, also following* [27]*, consider the subspace $X_2 = \{v \in L^2(\Omega) : \int_\Omega v = 0\}$ of both $L^2(\Omega)$ and $G(\Omega)$ which coincides with the space $\{v = \operatorname{div}(\vec{g}) : \vec{g} \cdot \vec{n} = 0\}$. However, the minimizers given by the $(BV(\Omega), G(\Omega))$ model will be different from the minimizers given by the ROF model* [34].

We would like now to introduce the space F (proposed earlier in [25] for the Navier–Stokes equations and in [27] as a suitable space for modeling textures instead of the space G). Let us work for a moment on the entire space $\mathbb{R}^2$ (assuming, for instance, that the data f is extended by zero or by reflection outside the open rectangle Ω).

DEFINITION 1.6 (John–Nirenberg space of bounded mean oscillation). *Let $f \in L^1_{loc}(\mathbb{R}^2)$. We say that f belongs to $BMO(\mathbb{R}^2)$ if the inequality*

$$\frac{1}{|Q|}\int_Q |f - f_Q| \le A$$

holds for all squares Q. (It is sufficient to consider squares with sides parallel with the axis.) Here $f_Q = |Q|^{-1}\int_Q f(x,y)$ denotes the mean value of f over the square Q. The smallest such A is chosen to be the norm of f in $BMO(\mathbb{R}^2)$, denoted by $\|f\|_{BMO(\mathbb{R}^2)}$, i.e.,

$$\|f\|_{BMO(\mathbb{R}^2)} = \sup_{Q=square} \frac{1}{|Q|}\int_Q |f - f_Q|. \tag{3}$$

Often in harmonic analysis, the Hardy space $H^1(\mathbb{R}^2)$ [36] is preferred instead of $L^1(\mathbb{R}^2)$ (with $H^1(\mathbb{R}^2) \subset L^1(\mathbb{R}^2)$), because $H^1(\mathbb{R}^2)$ has a predual which is the space $VMO(\mathbb{R}^2)$ (vanishing mean oscillation space [36]) while $L^1(\mathbb{R}^2)$ does not. When this substitution is applied, $L^\infty(\mathbb{R}^2) = (L^1(\mathbb{R}^2))'$ is substituted by $BMO(\mathbb{R}^2) = (H^1(\mathbb{R}^2))'$. Therefore, we are led to consider in a similar way the space F of generalized functions defined as (Meyer [27] and Koch and Tataru [25]) follows.

DEFINITION 1.7. *Let F consist of generalized functions T which can be written as*

$$T = \operatorname{div}(\vec{g}), \quad \vec{g} = (g_1, g_2) \in BMO(\mathbb{R}^2, \mathbb{R}^2).$$

Define $\|.\|_F$ on F by

$$\|T\|_F = \inf\big\{ \big(\|g_1\|_{BMO(\mathbb{R}^2)} + \|g_2\|_{BMO(\mathbb{R}^2)}\big) : T = \operatorname{div}(\vec{g}), \\ \vec{g} = (g_1, g_2) \in BMO(\mathbb{R}^2, \mathbb{R}^2)\big\}.$$

Similar with the case of the space G, this space F can also be identified with the dual of the $H^1(\mathbb{R}^2)$-Sobolev space $I_1(H^1)$ (with I_1 the Riesz potential) (see [37]) or sometimes denoted by $F^1_{1,2}(\mathbb{R}^2)$ (see [42]). Meyer suggests that the space F can also better model the oscillatory component v in the $u+v$ decomposition model than the L^2 space, and we will show this statement in this paper. In the rest of the paper, we will work with local versions $F(\Omega)$, since $BMO(\Omega)$ is well defined also on bounded domains (even if the Hardy space H^1 is not well suited on bounded domains).

DEFINITION 1.8. *Let $F(\Omega)$ consist of generalized functions T in $\mathcal{D}'(\Omega)$ which can be written*

$$T = \operatorname{div}(\vec{g}), \quad \vec{g} = (g_1, g_2) \in BMO(\Omega, \mathbb{R}^2),$$

equipped with the norm

$$\|T\|_{F(\Omega)} = \inf\big\{ \big(\|g_1\|_{BMO(\Omega)} + \|g_2\|_{BMO(\Omega)}\big) : T = \operatorname{div}(\vec{g}), \\ \vec{g} = (g_1, g_2) \in BMO(\Omega, \mathbb{R}^2)\big\}.$$

The next two examples show why the choice of $X_2 = L^2$ or $X_2 = L^1$ does not always model oscillatory functions very well, and the proposed models, obtained by substituting $\|v\|_{L^2}^2$ by $\|v\|_F$ or $\|v\|_G$ in the ROF model, give better decompositions.

EXAMPLE 1. *Let $a > 0$, $n > 0$ be fixed, and let φ be a smooth function defined on $\mathbb{R}$ such that*

$$\varphi(x) = \begin{cases} a & \text{if } |x| < n, \\ 0 & \text{if } |x| > n+1, \end{cases}$$

and φ is increasing on $(-\infty, 0]$ and decreasing on $[0, \infty)$. Let $m > 0$, $f(x) = \frac{1}{m}\varphi'(x)\sin(mx) + \varphi(x)\cos(mx)$, and $g(x) = \frac{\varphi(x)}{m}\sin(mx) + c$; then $f = g'$.

(i) *We have $\|f\|_G = \frac{a}{m}$. Note that $\|f\|_F < 2\|f\|_G \to 0$ as $m \to \infty$. (The inequality $\|f\|_F < 2\|f\|_G$ will be seen later.)*

(ii) $\|f\|_{L^2}^2 \geq 2a^2 \int_0^n |\cos(mx)|^2 \, dx = a^2(n + \frac{1}{2m}\sin(2mn)) \to a^2 n > 0$ *as* $m \to \infty$.

(iii) *Let $N = \lfloor \frac{mn}{2\pi} \rfloor$ be the number of complete periods of $\cos(mx)$ in the interval $[0, n]$. We may assume $N \geq 1$. Then*

$$\|f\|_{L^1} \geq 2a \int_0^n |\cos(mx)| \, dx \geq 8aN \int_0^{\frac{\pi}{2m}} cos(mx) \, dx = \frac{8aN}{m} \approx 4an/\pi.$$

Therefore, an oscillatory function has small G and F norms which do not depend on the domain $\Omega = (-n, n)$ and approach 0 as the frequency of oscillations increases but with important, not so small, L^2 and L^1 norms.

EXAMPLE 2. *Let $D = D(0, R) \subset \mathbb{R}^2$ be a disk centered at the origin with radius R. For some $\alpha > \epsilon > 0$, consider $f = \alpha\chi_D$, $u_\epsilon = (\alpha - \epsilon)\chi_D$, and $v_\epsilon = f - u_\epsilon = \epsilon\chi_D$. If we evaluate and compare the ROF energy for two candidate solutions, $(u = f, v = 0)$ and $(u_\epsilon, v_\epsilon = f - u_\epsilon)$, we would like to have for this f, for any ϵ,*

$$|f|_{BV} = \mathcal{J}(f, 0) \leq |u_\epsilon|_{BV} + \lambda\|v_\epsilon\|_{L^2}^2 \\ \text{or } 2\pi R\alpha \leq 2\pi R\alpha + \epsilon\pi R(\epsilon R\lambda - 2)$$

if and only if $\lambda \geq \frac{2}{R\epsilon}$ for all ϵ, i.e., when $\lambda = \infty$. However, if $\|.\|_{L^2}^2$ is replaced by $\|.\|_F$, then we have, for any ϵ,

$$|f|_{BV} = \mathcal{J}(f, 0) \leq |u_\epsilon|_{BV} + \lambda\|v_\epsilon\|_F \leq |u_\epsilon|_{BV} + 2\lambda\|v_\epsilon\|_G \\ \Leftrightarrow 2\pi\alpha R \leq 2\pi R\alpha + \epsilon R(\lambda - 2\pi)$$

if and only if $\lambda > 2\pi$ which does not depend on ϵ and R. (To obtain $\|v_\epsilon\|_G = \epsilon\frac{R}{2}$, we use [27, *Lemma* 6, *page* 36].*)*

This shows a limitation of the ROF model: if $f = \alpha\chi_D$ (an image free of noise, piecewise-constant, and with smooth discontinuity set of finite length), then the minimizer u_λ cannot be f for any finite λ. Related remarks have been made in [27], [38], [39], [5]. Moreover, suppose we decompose f by the energy minimization

$$\inf_{(u,v)} \{\mathcal{K}(u,v) = |u|_{BV} + \lambda\|v\|_{X_2}^p : f = u + v, p > 0\},$$

where $\|\cdot\|_{X_2}$ is a norm or a quasi norm, and such that $\|\chi_D\|_{X_2} \neq 0$. If we start with $f = \alpha\chi_D$, then the recovered image should be f for some finite $\lambda > C$. We have the following: $\mathcal{K}(f,0) \leq \mathcal{K}(u_\epsilon, v_\epsilon)$ if and only if $p \leq 1$. Therefore, for any $p > 1$, we cannot obtain $u = f$ for any finite value of λ. We refer the reader to Cheon et al. [15], Chan and Esedoglu [11], and Chan, Esedoglu, and Nikolova [12] for the $(BV(\Omega), L^1(\Omega))$ version of the ROF model in two dimensions in the continuous setting, which are also improvements over the original ROF model. (In [15], the authors have also considered the case when in the ROF model the fidelity term $\|f-u\|_{L^2(\Omega)}^2$ has been replaced by $\|f-u\|_{L^2(\Omega)}$; this choice gives very good reconstruction results in practice.)

From such motivations, Meyer [27] proposed a decomposition of f, with $X_1 = BV(\Omega)$, via

$$\inf\{\mathcal{K}(u,v) = |u|_{BV(\Omega)} + \lambda\|v\|_{X_2}\},$$

where the infimum is taken over $u \in BV(\Omega)$ and $v \in X_2$, such that $f = u + v$. Here $(X_2, \|\cdot\|_{X_2})$ is either $(G(\Omega), \|.\|_{G(\Omega)})$ or $(F(\Omega), \|.\|_{F(\Omega)})$. However, these minimization models cannot be directly solved in practice: there is no standard calculation of the associated Euler–Lagrange equation, as it is for the ROF model which can be solved easily by finite differences.

In [44], [45], Vese and Osher proposed a method to overcome the difficulty of computing $\|.\|_G$. This has been done by the energy minimization problem

$$\begin{aligned}\inf_{u,g_1,g_2} \Big\{ &\mathcal{G}_p(u,g_1,g_2) = |u|_{BV(\Omega)} + \mu\|f - u - \partial_x g_1 - \partial_y g_2\|_{L^2(\Omega)}^2 \\ &+ \lambda\|\sqrt{g_1^2 + g_2^2}\|_{L^p(\Omega)}\Big\}.\end{aligned} \tag{4}$$

By this model, f is decomposed into $u+v+w$, and as $\mu \to \infty$ and $p \to \infty$, the model approaches Meyer's (BV, G) model. The space $G = W^{-1,\infty}(\Omega)$ is approximated by $W^{-1,p}(\Omega)$, with $p < \infty$. (When $p = 2$, v belongs to the dual of the Sobolev space $H_0^1(\Omega)$.)

In [33], Osher, Solé, and Vese proposed a simplified approximated method corresponding to the case $p = 2$. Let $\vec{g} = \nabla P + \vec{Q}$ be the unique Hodge decomposition of $\vec{g} \in L^\infty(\Omega, \mathbb{R}^2)$. Using $f - u = v = \operatorname{div}(\vec{g}) = \triangle P$, i.e., $P = \triangle^{-1}(f-u)$, Meyer's (BV, G) model is then approximated by

$$\inf_u \left\{ \mathcal{G}_2(u) = \int_\Omega |\nabla u| + \lambda \int_\Omega |\nabla(\triangle^{-1}(f-u))|^2 \right\}.$$

This model gives an exact decomposition $f = u + v$, with $u \in BV(\Omega)$ and $v \in W^{-1,2}(\Omega) = (H_0^1(\Omega))'$, and the minimization problem has been solved using a fourth-order nonlinear PDE.

In the present paper, we propose a new method to approximate Meyer's (BV, F) model. We also introduce an equivalent definition of the BMO norm, using an open

set formulation, which is easily formulated and computed using curve evolution technique. As a by-product, we also propose a new method for solving the (BV, G) model, different from the one proposed in [7], [6].

As we have mentioned, Mumford and Gidas [29] show that natural images, as samples from scale-invariant probability distributions, cannot be modeled by functions but instead by generalized functions, i.e., distributions in $\mathcal{D}'(\Omega)$.

Other related models for image decomposition into cartoon and texture have been proposed recently. We mention Daubechies and Teschke [16] and Starck, Elad, and Donoho [35] for variational and wavelets approaches.

In particular, we refer the reader to Aujol et al. [7] and Aubert and Aujol [6] for more properties of the space G both in theory and practice and to another approximation of the Meyer's (BV, G) model on bounded domains. We also refer the reader to Aujol and Chambolle [8] for properties of norms that are dual to negative Sobolev and Besov norms. Our theoretical framework extends some of the results presented in Aubert and Aujol [6] for the space $G(\Omega)$ to the case of the space $F(\Omega)$.

Other related works are by Esedoglu and Osher [19], Osher and Scherzer [32], Obereder, Osher, and Scherzer [31], and Goldfarb and Yin [22], among others.

We believe that the case (BV, F) has not been considered in theory or in practice previously in image analysis; therefore, our contribution is new also from this point of view.

The theoretical work of Koch and Tataru [25] (mentioned by Meyer in [27]) uses the space $\mathrm{div}(BMO(\mathbb{R}^2))$ for solutions of the Navier–Stokes equations. Finally, in [9], Bourgain and Brezis analyze the equation $f = \mathrm{div}(\vec{y})$ in some limiting cases, and applications of such results can be found in Aubert and Aujol [6] for the analysis of the space $G(\Omega)$.

2. Definitions and properties of the *BMO* space. Here we would like to review the definitions and some basic properties of the space BMO. We refer the reader to *"Harmonic Analysis"* by Stein [36] and also to [41], [24], and [20].

Let Ω be an open and bounded subset of $\mathbb{R}^n$. For planar images, we may assume that $\Omega = (0,1) \times (0,1) \subset \mathbb{R}^2$. To simplify the notation, we will often write BV, G, F, ..., instead of $BV(\Omega)$, $G(\Omega)$, $F(\Omega)$, etc.

DEFINITION 2.1. *Let $f \in L^1_{loc}$. We say that f belongs to BMO^β if the inequality*

$$\frac{1}{|O|} \int_O |f - f_O| \, dx \le A_1$$

holds for the family $\mathcal{F}^\beta$ of open sets $O \subset \Omega$ such that there exist cubes Q_1, and Q_2 with $Q_1 \subset O \subset Q_2 \subset \Omega$, and $\frac{|Q_2|}{|Q_1|} \le \beta$; here $1 \le \beta < \infty$ is a constant, and $f_O = |O|^{-1} \int_O f \, dx$. The smallest such A_1 is chosen to be the norm of f in BMO^β, denoted by $\|f\|_{BMO^\beta}$, i.e.,

$$\|f\|_{BMO^\beta} = \sup_{O \in \mathcal{F}^\beta} \frac{1}{|O|} \int_O |f - f_O| \, dx. \tag{5}$$

DEFINITION 2.2. *Let $f \in L^1_{loc}$. We say that f belongs to BMO if the inequality*

$$\frac{1}{|Q|} \int_Q |f - f_Q| \, dx \le A_2$$

holds for all cubes $Q \subset \Omega$ with sides parallel with the axes. The smallest such A_2 is chosen to be the norm of f in BMO, denoted by $\|f\|_{BMO}$, i.e.,

$$\|f\|_{BMO} = \sup_{Q\subset\Omega} \frac{1}{|Q|} \int_Q |f - f_Q| \, dx. \tag{6}$$

$\|.\|_{BMO}$ is a seminorm vanishing on constant functions. If we identify functions in BMO which are different a.e. by a constant, then BMO becomes a Banach space. We obtain an equivalent norm if the family of cubes is replaced by the family of balls. Moreover, as mentioned in [36],

$$\|f\|_{BMO_p} = \left[\sup_{Q\subset\Omega} \frac{1}{|Q|} \int_Q |f - f_Q|^p \, dx\right]^{\frac{1}{p}} \tag{7}$$

gives an equivalent BMO norm for $p \geq 1$. Here we will consider the cases $p = 1$ and $p = 2$.

Definition 2.3. *A dyadic cube is a cube of the special form*

$$Q = \{k_j 2^{-m} < x_j < (k_j + 1)2^{-m}; 1 \leq j \leq n\}, \tag{8}$$

where m and k_j, $1 \leq j \leq n$, are integers. We say f has bounded dyadic mean oscillation, $f \in BMO_d$, if

$$\|f\|_{BMO_d} = \sup_{Q\subset\Omega \ dyadic} \frac{1}{|Q|} \int_Q |f - f_Q| \, dx < \infty.$$

Let $T_\alpha f(x) = f(x - \alpha)$. We say $f \in BMO_{d,\alpha}$ if $T_\alpha f \in BMO_d$.

Let $A = \left\{\alpha = (\alpha_1, \ldots, \alpha_n) : \alpha_i \in \left\{0, \frac{1}{3}\right\}\right\}$. Note $|A| = 2^n$.

Lemma 2.4 (see [21] and [26]). *$f \in BMO$ if and only if $f \in BMO_{d,\alpha}$ for all $\alpha \in A$. In fact, $\|f\|_{BMO} \leq 12 \max_{\alpha\in A} \left\{\|f\|_{BMO_{d,\alpha}}\right\}$.*

Remark 2. *Let $BMO_{\mathcal{D}} = \cap_{\alpha\in A} BMO_{d,\alpha}$ with*

$$\|.\|_{BMO_{\mathcal{D}}} = \max_{\alpha\in A}\{\|.\|_{BMO_{d,\alpha}}\}.$$

Then the above lemma shows that $BMO = BMO_{\mathcal{D}}$ with equivalent norms.

Lemma 2.5. *Let $f \in L^1_{loc}$, and let $c \in \mathbb{R}$; then*

$$\int_O |f - f_O| \, dx \leq 2 \int_O |f - c| \, dx \tag{9}$$

for all $O \subset \Omega$.

Proof.

$$\begin{aligned} |f - f_O| &\leq |f - c| + |c - f_O| = |f - c| + |c - |O|^{-1} \int_O f \, dx| \\ &\leq |f - c| + |O|^{-1} \int |c - f| \, dx. \end{aligned}$$

Integrating both sides over O, we obtain (9). Moreover,

$$\inf_{c\in\mathbb{R}} \int_O |f - c| \, dx \leq \int_O |f - f_O| \, dx \leq 2 \inf_{c\in\mathbb{R}} \int_O |f - c| \, dx. \qquad \square$$

LEMMA 2.6. *$\|.\|_{BMO^\beta}$ and $\|.\|_{BMO}$ are equivalent for any $1 \le \beta < \infty$.*

Proof. It suffices to show that there exist constants c_1 and c_2 greater than 0 such that, for all $f \in BMO$,

$$c_1\|f\|_{BMO} \le \|f\|_{BMO^\beta} \le c_2\|f\|_{BMO}. \tag{10}$$

It is clear that the first inequality in (10) holds with $c_1 = 1$. It remains to show $\|f\|_{BMO^\beta} \le c_2\|f\|_{BMO}$ for some $c_2 > 0$.

Let $O \in \mathcal{F}^\beta$. There exist Q_1 and Q_2 in Ω such that $Q_1 \subset O \subset Q_2$, and $\frac{|Q_2|}{|Q_1|} \le \beta$. We have

$$\begin{aligned}\frac{1}{|O|}\int_O |f - f_O|\,dx &\le 2\inf_{c\in\mathbb{R}} \frac{1}{|O|}\int_O |f-c|\,dx \le 2\inf_{c\in\mathbb{R}} \frac{1}{|Q_1|}\int_{Q_2} |f-c|\,dx\\ &\le 2\beta \inf_{c\in\mathbb{R}} \frac{1}{|Q_2|}\int_{Q_2} |f-c|\,dx \le 2\beta\left(\frac{1}{|Q_2|}\int_{Q_2}|f-f_{Q_2}|\,dx\right)\\ &\le 2\beta \sup_{Q\subset\Omega}\frac{1}{|Q|}\int_Q |f-f_Q|\,dx = \|f\|_{BMO}.\end{aligned}$$

Taking the supremum over all $O \in \mathcal{F}^\beta$, we have $\|f\|_{BMO^\beta} \le c_2\|f\|_{BMO}$ with $c_2 = 2\beta$. ∎

The next simple property shows that functions in BMO are scale invariant. For simplicity, assume here that $\Omega = \mathbb{R}^n$.

LEMMA 2.7. *$f(x)$ and $f(\alpha x)$ have the same norm in BMO for all $\alpha > 0$.*

Proof. Using a change of variable by letting $y = \alpha x$, we have

$$\begin{aligned}\frac{1}{|Q|}\int_Q \left|f(\alpha x) - \frac{\int_Q f(\alpha x)\,dx}{|Q|}\right| dx &= \frac{1}{|Q|}\int_{\alpha Q}\left|f(y) - \frac{\int_{\alpha Q} f(y)\frac{1}{\alpha^n}dy}{|Q|}\right|\frac{1}{\alpha^n}dy\\ &= \frac{1}{|\alpha Q|}\int_{\alpha Q}\left|f(y) - \frac{\int_{\alpha Q} f(y)dy}{|\alpha Q|}\right|dy\\ &= \frac{1}{|Q'|}\int_{Q'}\left|f(y) - \frac{\int_{Q'} f(y)dy}{|Q'|}\right|dy,\end{aligned} \tag{11}$$

where $Q' = \alpha Q = \{\alpha x : x \in Q\}$; it is clear that

$$\sup_{Q\subset R^n}\frac{1}{|Q|}\int_Q \left|f(x) - \frac{\int_Q f(x)\,dx}{|Q|}\right| dx = \sup_{Q'\subset R^n}\frac{1}{|Q'|}\int_{Q'}\left|f(y) - \frac{\int_{Q'} f(y)dy}{|Q'|}\right|dy. \qquad ∎$$

Note that the norms on G and F are not scale invariant but satisfy the scaling relation $\|f(\alpha\cdot)\|_{F,G} = \frac{1}{\alpha}\|f(\cdot)\|_{F,G}$.

REMARK 3. *$L^\infty \subset BMO^\beta = BMO$, for $1 \le \beta < \infty$, with $\|f\|_{BMO} \le 2\|f\|_\infty$. Note that if in the definition of BMO^β we do not impose any bound on β, i.e., we allow $\beta = \infty$, then BMO^β approaches the space L^∞: the norm is attained at a union of very small regions, such that f has largest or smallest values inside these regions. BMO also contains unbounded functions; indeed, $ln(|P|) \in BMO$ for any polynomial P.*

Some additional properties are as follows. $|f|$ is in BMO whenever f is, since $\|f| - |f|_Q| \le |f - f_Q|$. However, if $|f|$ is in BMO, then f is not necessarily in BMO.

For example in $\mathbb{R}$, $f(x) = sign(x)ln(|x|)$ is not in $BMO(\mathbb{R})$, while $|f(x)| = ln(|x|)$ is. Indeed, for $\alpha > 0$, $f(\alpha x) = sign(x)ln(\alpha) + sign(x)ln(|x|)$, and if $I = [-1, 1]$, then

$$\frac{1}{|\alpha I|}\int_{\alpha I}\left|f(y) - \frac{\int_{\alpha I} f(y)dy}{|\alpha I|}\right| dy \to \infty \text{ as } \alpha \to \infty.$$

3. Numerical computation of the BMO norm. In this section, we introduce and discuss several new methods for computing or approximating the BMO norm of a given function f in two dimensions, using the equivalent definitions introduced in the previous section.

3.1. Computing the BMO norm using the open set formulation. Let ϕ be a Lipschitz-continuous function that defines ∂O implicitly, i.e., $O = \{x \in \Omega : \phi(x) > 0\}$, and denote by $H(\phi)$ the Heaviside function

$$H(r) = \begin{cases} 1 & \text{if } r \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Using the variational level set formulation, as in [46], [13], [14], we define

$$\mathcal{G}(\phi) = \frac{1}{\int_\Omega H(\phi)\ dx}\int_\Omega \left|g - \frac{\int_\Omega gH(\phi)\ dx}{\int_\Omega H(\phi)\ dx}\right| H(\phi)\ dx. \tag{12}$$

Similarly, we also have an equivalent $\|.\|_{BMO^\beta}$ norm using $p = 2$:

$$\mathcal{F}(\phi) = \left[\frac{1}{\int_\Omega H(\phi)\ dx}\int_\Omega \left|g - \frac{\int_\Omega gH(\phi)\ dx}{\int_\Omega H(\phi)\ dx}\right|^2 H(\phi)\ dx\right]^{\frac{1}{2}}. \tag{13}$$

REMARK 4. *We will incorporate in* (12) *the perimeter of the unknown open set O as given by $\int_\Omega |\nabla H(\phi)|$ in order to ensure that O remains "bulky"; i.e., we do not allow O to break into very small pieces. We do not exactly impose in this way the required constraint from Definition* 2.1*, but it is a way of keeping the ratio between $|O|$ and ∂O bounded. This was kindly suggested to us by Jean-Michel Morel.*

If ϕ solves $\sup_\phi\{\mathcal{G}_{new}(\phi) = \mathcal{G}(\phi) - \lambda\int_\Omega |\nabla H(\phi)|\ dx\}$ for some parameter $\lambda > 0$, then $\|g\|_{BMO^\beta}$ is well approximated by $\mathcal{G}(\phi)$.

Let $H_\epsilon(\phi)$ be a smoother function approximating $H(\phi)$ as $\epsilon \to 0$. Using the notation $|O| = \int_\Omega H_\epsilon(\phi)\ dx$, we obtain

$$\begin{aligned}\frac{\partial \mathcal{G}_{new}}{\partial \phi} = \Bigg[&- \frac{1}{|O|^2}(g - g_O)\int_\Omega \frac{g - g_O}{|g - g_O|}H_\epsilon(\phi)\ dx + \frac{1}{|O|}|g - g_O| - \frac{\mathcal{G}(\phi)}{|O|} \\ &+ \lambda \text{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right)\Bigg]\delta_\epsilon(\phi),\end{aligned}$$

where $\delta_\epsilon = H'_\epsilon$. By introducing an artificial time, we then solve

$$\phi_t = \frac{\partial \mathcal{G}_{new}}{\partial \phi} \text{ in } \Omega \quad \text{and} \quad \frac{\nabla u}{|\nabla u|}\cdot \vec{n} \text{ on } \partial\Omega.$$

However, in practice, since our approximation $\delta_\epsilon(\phi) > 0$ for any ϕ as in [13], we neglect this factor $\delta_\epsilon(\phi)$, and we solve the equation (to obtain faster results)

$$\begin{aligned}\phi_t = &- \frac{1}{|O|^2}(g - g_O)\int_\Omega \frac{g - g_O}{|g - g_O|}H(\phi)\ dx + \frac{1}{|O|}|g - g_O| - \frac{\mathcal{G}(\phi)}{|O|} \\ &+ \lambda \text{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right).\end{aligned}$$

In our numerical calculations, we always have that $Q_2(O) = \Omega$ (therefore, $|Q_2|$ is bounded independent of O), while the existence of $Q_1(O)$, with size that does not become too small, is ensured by the additional length term.

We do not guarantee that we compute a global maximum of the energy. However, the numerical experiments (using piecewise-constant images) show that we obtain good and stable approximations to the exact solution, as illustrated in the section of experimental results.

3.2. Computing the *BMO* norm using the square formulation. Note that, in two dimensions, the set $Q' = \{|x|+|y| < r\}$ is a square centered at the origin, with side length $l(Q') = \sqrt{2}r$. The corners of Q' are at the vertices $(r, 0)$, $(0, r)$, $(-r, 0)$, $(0, -r)$. Therefore, to have the sides of Q' of length r and parallel to the axis, we need to rotate $\frac{1}{\sqrt{2}}Q'$ by an angle of $\frac{\pi}{4}$, i.e., by applying the matrix

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

to Q'. In other words, $Q = \{|(x - x_0) - (y - y_0)| + |(x - x_0) + (y - y_0)| < r\}$ is the square centered at (x_0, y_0), with sides parallel to the axis and $l(Q) = r$.

Let $\phi(x, y) = r-(|(x-x_0)-(y-y_0)|+|(x-x_0)+(y-y_0)|)$ be the signed distance function to ∂Q, and let $H(\phi)$ be the Heaviside function. Define

$$\mathcal{G}(Q) = \mathcal{G}(r, x_0, y_0) = \frac{1}{\int_\Omega H(\phi)} \int_\Omega \left| g - \frac{\int_\Omega gH(\phi)}{\int_\Omega H(\phi)} \right| H(\phi). \tag{14}$$

Then

$$\|g\|_{BMO} = \sup_{(r,x_0,y_0)} \mathcal{G}(r, x_0, y_0).$$

We also have an equivalent $\|.\|_{BMO}$ norm with $p = 2$:

$$\mathcal{F}(Q) = \mathcal{F}(r, x_0, y_0) = \left[\frac{1}{\int_\Omega H(\phi)} \int_\Omega \left| g - \frac{\int_\Omega gH(\phi)}{\int_\Omega H(\phi)} \right|^2 H(\phi) \right]^{\frac{1}{2}}. \tag{15}$$

Let H_ϵ be a smooth approximation of H, and let $\delta_\epsilon = H'_\epsilon$. We have

$$\begin{aligned}
\frac{\partial \mathcal{G}}{\partial r} =& \frac{1}{\int_\Omega H_\epsilon(\phi)} \Big[- \mathcal{G} \int_\Omega \frac{\partial H_\epsilon(\phi)}{\partial r} + \int_\Omega |g - g_Q| \frac{\partial H_\epsilon(\phi)}{\partial r} \\
& - \frac{1}{\int_\Omega H_\epsilon(\phi)} \int_\Omega \frac{g - g_Q}{|g - g_Q|} H_\epsilon(\phi) \int_\Omega (g - g_Q) \frac{\partial H_\epsilon(\phi)}{\partial r} \Big], \\
\frac{\partial \mathcal{G}}{\partial x_0} =& \frac{1}{\int_\Omega H_\epsilon(\phi)} \Big[- \mathcal{G} \int_\Omega \frac{\partial H_\epsilon(\phi)}{\partial x_0} + \int_\Omega |g - g_Q| \frac{\partial H_\epsilon(\phi)}{\partial x_0} \\
& - \frac{1}{\int_\Omega H_\epsilon(\phi)} \int_\Omega \frac{g - g_Q}{|g - g_Q|} H_\epsilon(\phi) \int_\Omega (g - g_Q) \frac{\partial H_\epsilon(\phi)}{\partial x_0} \Big],
\end{aligned}$$

where $\frac{\partial H_\epsilon(\phi)}{\partial r} = \delta_\epsilon(\phi)$, and $\frac{\partial H_\epsilon(\phi)}{\partial x_0} = \big[\frac{(x-x_0)-(y-y_0)}{|(x-x_0)-(y-y_0)|} + \frac{(x-x_0)+(y-y_0)}{|(x-x_0)+(y-y_0)|}\big]\delta_\epsilon(\phi)$. Similarly,

$$\begin{aligned}
\frac{\partial \mathcal{G}}{\partial y_0} =& \frac{1}{\int_\Omega H_\epsilon(\phi)} \Big[- \mathcal{G} \int_\Omega \frac{\partial H_\epsilon(\phi)}{\partial y_0} + \int_\Omega |g - g_Q| \frac{\partial H_\epsilon(\phi)}{\partial y_0} \\
& - \frac{1}{\int_\Omega H_\epsilon(\phi)} \int_\Omega \frac{g - g_Q}{|g - g_Q|} H_\epsilon(\phi) \int_\Omega (g - g_Q) \frac{\partial H_\epsilon(\phi)}{\partial y_0} \Big],
\end{aligned}$$

where $\frac{\partial H_\epsilon(\phi)}{\partial y_0} = \big[- \frac{(x-x_0)-(y-y_0)}{|(x-x_0)-(y-y_0)|} + \frac{(x-x_0)+(y-y_0)}{|(x-x_0)+(y-y_0)|}\big]\delta_\epsilon(\phi)$.

By introducing an artificial time, we will solve the equations

$$\frac{\partial r}{\partial t} = \frac{\partial \mathcal{G}}{\partial r}, \quad \frac{\partial x_0}{\partial t} = \frac{\partial \mathcal{G}}{\partial x_0}, \quad \frac{\partial y_0}{\partial t} = \frac{\partial \mathcal{G}}{\partial y_0}.$$

Again, we do not show that this method converges to a global maximum of the energy. However, in the experimental results (using piecewise-constant images), we have obtained the correct answer when we know the exact solution.

If we would work with disks instead of squares, then we could have

$$\phi(x, y) = r^2 - (x - x_0)^2 - (y - y_0)^2,$$

which is differentiable everywhere.

3.3. Exact computation of the BMO norm. As kindly suggested by Jean-Michel Morel, we have also implemented an exact evaluation of the BMO norm using the square formulation. This is computationally more expensive but still can be made relatively fast by using FFT. In addition, it produces very accurate results. The procedure is as follows.

1. Fix a list of growing scales $\sigma = 2, 4, 8, 16, \ldots$.
2. For each σ, consider the function $k_{\sigma,(x,y)} = \frac{1}{\sigma^2}\chi_{Q_{\sigma,(x,y)}}$, where $Q_{\sigma,(x,y)}$ is the square centered at (x, y) having length σ.
3. Convolve f with $k_{\sigma,(0,0)}$, called f_σ.
4. For each (x, y), compute $osc(x, y) = \int |f - f_\sigma| k_{\sigma,(x,y)}$.
5. Take the sup of $osc(x, y)$ which yields a value of $\|f\|_{BMO}$ at scale σ, denoted $\|f\|_{BMO,\sigma}$.
6. Compute the maximal value of $\|f\|_{BMO,\sigma}$ for all σ's.

If one takes $\sigma = 2, 3, 4, 5, \ldots, N$, where N is the size of the image f, one is ensured to get the exact value of $\|f\|_{BMO}$ in a discrete framework. In this way, one obtains the global maximum and the square(s) where it is attained. This allows one to compare with the maximization procedure and see whether or not it yields the same value.

Finally, we have used the above procedure to evaluate the necessary expression over the dyadic squares and the additional translations. This gives an accurate method too, and it is faster since we have fewer squares to consider.

4. A (BV, F) image decomposition model. Recall Meyer's model, which decomposes f into $u + v$, by the variational problem

$$\inf\{\mathcal{E}(u, v) = |u|_{BV} + \lambda \|v\|_F\}, \tag{16}$$

where the infimum is taken over $u \in BV$ and $v \in F$, such that $f = u + v$.

The space F is defined as

$$F = \{v = \operatorname{div}(\vec{g}) \text{ in } \mathcal{D}' : \vec{g} \in BMO(\Omega, \mathbb{R}^2)\}.$$

REMARK 5. *Given $f \in L^2$, there exists $u \in BV$, and $v \in L^2 \subset G \subset F$, such that $\mathcal{E}(u, v) < \infty$, $f = u + v$, and $\int f = \int u$.*

Indeed, recall the ROF model, which minimizes the functional

$$\mathcal{J}(u, v) = |u|_{BV} + \lambda' \|v\|_{L^2} \tag{17}$$

over the set of $u \in BV$, $v \in L^2$, such that $f = u + v$. Pick λ' so that $\|f\|_G > \frac{1}{2\lambda'}$. The existence of a minimizer, denoted (u, v), for the ROF model has been proved in [10]

and characterized in [27] *as being* $u \in BV$, $v \in G \subset F$, *with* $\|v\|_G = \frac{1}{2\lambda'}$. *Therefore,* $\mathcal{E}(u,v) < \infty$.

The next theorem shows the existence of minimizers for Meyer's (BV, F) model. We refer the reader to Aubert and Aujol [6] for a similar proof when $\|.\|_* = \|.\|_G$.

THEOREM 4.1. *Let* $f \in L^2$. *The minimization problem*

$$\inf_{(u,v)} \left\{ \mathcal{E}(u,v) = |u|_{BV} + \lambda \|v\|_F, \ \int_\Omega u = \int_\Omega f, \ f = u + v \right\} \tag{18}$$

has at least one solution $u \in BV$, $v = f - u \in F \cap L^2$.

Proof. We use the standard tool in calculus of variations. Let $\{(u_n, v_n)\}$ be a minimizing sequence. (From the previous remark, we know that the infimum of the energy is finite.) Then $f = u_n + v_n$ and $\int_\Omega u_n = \int_\Omega f$ for all $n \geq 0$. In addition, there is a constant C (that may change from line to line) such that

$$\begin{aligned} |u_n|_{BV} &\leq C, \\ \|v_n\|_F &\leq C \end{aligned}$$

uniformly.

By Poincare–Wirtinger inequality,

$$\left\| u_n - \int_\Omega u_n \right\|_{L^2} \leq C |u|_{BV},$$

and since $\int_\Omega u_n = \int_\Omega f$, for all n,

$$\begin{aligned} &\Rightarrow \|u_n\|_{L^2} \leq C \\ &\Rightarrow |u_n|_{BV} \leq C. \end{aligned}$$

Then $\|u_n\|_{L^1} \leq C$ since Ω is bounded; furthermore, $\|u_n\|_{BV} = \|u_n\|_{L^1} + |u_n|_{BV} \leq C$. Therefore, there exists $u \in BV$ and a subsequence (still denoted by u_n), such that u_n converges to u in the BV-weak* topology. In particular, u_n converges to u strongly in L^1, and by the lower semicontinuity of the total variation, $|u|_{BV} \leq \liminf_{n\to\infty} |u_n|_{BV}$.

As for the subsequence v_n, we have $v_n = \operatorname{div}(\vec{g}_n)$ in $\mathcal{D}'$ and $\|v_n\|_F = \|g_{1,n}\|_{BMO} + \|g_{2,n}\|_{BMO} \leq C$, and we obtain that $g_{i,n} \to g_i$ in *BMO-weak**. We also have, for all $\phi \in \mathcal{D}$, $\int_\Omega v_n \phi = -\int_\Omega \vec{g}_{i,n} \cdot \nabla\phi \to -\int_\Omega \vec{g}_i \cdot \nabla\phi$. Therefore, $v = \operatorname{div}(\vec{g})$ in $\mathcal{D}'$.

Since $\|v_n\|_{L^2} \leq C$, up to a subsequence, $v_n \to v$ weakly in L^2, and we have $v = \operatorname{div}(\vec{g})$ a.e. As $v_n = f - u_n$ and u_n converges to u weakly in L^2, we also obtain $v = f - u$ a.e.

By weak* lower semicontinuity, it follows that

$$\begin{aligned} \|v\|_F \leq \|g_1\|_{BMO} + \|g_2\|_{BMO} &\leq \|g_{1,n}\|_{BMO} + \|g_{2,n}\|_{BMO} = \|v_n\|_F, \\ |u|_{BV} &\leq \liminf |u_n|_{BV}, \\ \|v\|_F &\leq \liminf \|v_n\|_F. \end{aligned}$$

Therefore, $\mathcal{E}(u,v) \leq \liminf_{n\to\infty} \mathcal{E}(u_n, v_n)$, and we obtain existence of minimizers. ∎

5. Approximating the (BV, F) decomposition model. Here we do not solve (16) directly, but we adapt the model [44] by adding a fidelity term into the energy. In this decomposition, $f \in L^2$ is decomposed into $u + v + w$, with $u \in BV$, $v \in F$, and a small residual $w \in L^2$.

The variational problem can be written as

$$\inf \left\{ \mathcal{E}_\mu(u,v) = |u|_{BV} + \mu \int_\Omega |f-u-v|^2 \ + \lambda \|v\|_F \right. \\ \left. : u \in BV, v \in F, \int_\Omega u \ = \int_\Omega f \ \right\}. \tag{19}$$

Taking $v = \operatorname{div}(\vec{g})$, we obtain an equivalent formulation in terms of u, g_1, and g_2:

$$\inf \left\{ \mathcal{E}_\mu(u,g_1,g_2) = |u|_{BV} + \mu \int_\Omega |f-u-\operatorname{div}(\vec{g})|^2 \right. \\ \left. + \lambda[\|g_1\|_{BMO} + \|g_2\|_{BMO}] \right\}, \tag{20}$$

where the infimum is taken over $g_i \in BMO$ and $u \in BV$ with

$$\int_\Omega u \ = \int_\Omega f \ . \tag{21}$$

The existence and uniqueness of a minimizer can be shown for the new model.

THEOREM 5.1. *Let $f \in L^2$. Then there exist $u \in BV$, and $v \in F \cap L^2$, such that (u,v) solves* (19) *or* (20). *If, in addition, $\int_\Omega f \neq 0$, then the minimizer is unique.*

Proof. Existence of minimizers: let (u_n, v_n) be a minimizing sequence of (19) or (20). We have

$$|u_n|_{BV} \leq C, \tag{22}$$

$$\|f - u_n - v_n\|_{L^2} \leq C, \tag{23}$$

$$\|v_n\|_F \leq C. \tag{24}$$

From the Poincaré inequality,

$$\left\| u_n - \int_\Omega u_n \right\|_{L^2} \leq C|u_n|_{BV}.$$

Since $\int_\Omega u_n = \int_\Omega f$, for all n, u_n is uniformly bounded in L^2. Since Ω is bounded, u_n is also uniformly bounded in L^1. Therefore,

$$\|u_n\|_{BV} \leq C. \tag{25}$$

Then there exists $u \in BV$, such that, up to a subsequence, u_n converges to u weak* in BV. By (23) and uniform boundedness of u_n in L^2, v_n is also uniformly bounded in L^2. Therefore, there exists $v \in L^2$ such that, up to a subsequence, v_n converges to v weakly in L^2.

As $\|v_n\|_F \leq C$, there exists $\vec{g}_n = (g_{1,n}, g_{2,n}) \in BMO(\Omega, \mathbb{R}^2)$, such that $v_n = \operatorname{div}(\vec{g}_n)$ in $\mathcal{D}'$, and $\|g_{i,n}\|_{BMO} \leq C$. Therefore, there exists $g_i \in BMO$, such that $g_{i,n}$ converges to g_i weak* in BMO, for $i = 1, 2$. Let $\vec{g} = (g_1, g_2)$.

To show $v = \operatorname{div}(\vec{g}) \in F$, let $\varphi \in \mathcal{D}$,

$$\int_\Omega v_n \varphi \ = \int_\Omega \operatorname{div}(\vec{g}_n)\varphi \ = -\int_\Omega \vec{g}_n \cdot \nabla\varphi.$$

Taking $n \to \infty$ (using weak L^2 topology and weak* $BMO(\Omega, \mathbb{R}^2)$ topology), we obtain

$$\int_\Omega v\varphi \ = -\int_\Omega \vec{g} \cdot \nabla\varphi \ = \int_\Omega \operatorname{div}(\vec{g})\varphi.$$

This implies $v = \operatorname{div}(\vec{g})$ in $\mathcal{D}'$ as a distribution. But since $v \in L^2$, $v = \operatorname{div}(\vec{g})$ a.e. Therefore, $v \in F \cap L^2$.

By weak and weak* lower semicontinuity, it follows that

$$\begin{gathered} |u|_{BV} \le \liminf |u_n|_{BV}, \\ \|f - u - v\|_{L^2} \le \liminf \|f - u_n - v_n\|_{L^2}, \text{ and} \\ \|v\|_F \le \liminf \|v_n\|_F. \end{gathered}$$

Therefore, $\mathcal{E}_\mu(u, v) \le \mathcal{E}_\mu(u_n, v_n)$, and (u, v) is a minimizer for (19).

Uniqueness of minimizers: denote by $(\hat{u}, \hat{v})$ a minimizer of the energy. Then $\int_\Omega \hat{u} = \int_\Omega f$.

The energy to be minimized is strictly convex, as the sum of two convex functions $(|u|_{BV} + \|v\|_F)$ and of a strictly convex function $\|f - (u + v)\|^2_{L^2}$, except in the direction $(u, -u)$ (as in [7], [6]). Therefore, it suffices to check that if $(\hat{u}, \hat{v})$ is a minimizer, then $(\hat{u} + t\hat{u}, \hat{v} - t\hat{u})$ is not a minimizer for $t \neq 0$. Since $(\hat{u}, \hat{v})$ is a minimizer, then $\int_\Omega \hat{u} = \int_\Omega f$. Therefore, if $(\hat{u} + t\hat{u}, \hat{v} - t\hat{u})$ is a minimizer too, then $\int_\Omega (1+t)\hat{u} = (1+t)\int_\Omega \hat{u} = \int_\Omega f$. This is possible only if $t = 0$; therefore, we conclude the uniqueness. □

Again as in [7], [6], we can show that the approximated model (19) approaches Meyer's model (18) as $\mu \to \infty$. In other words, we have the following theorem.

THEOREM 5.2. *Assume $f \in L^2$ with $\int_\Omega f \neq 0$, and let us assume that problem* (18) *has a unique solution $(\hat{u}, \hat{v})$. Let us denote by (u_μ, v_μ) the unique solution of* (19)*. Then, as $\mu \to \infty$, $u_\mu + v_\mu \to f$ strongly in L^2, and (u_μ, v_μ) converges to some (u_0, v_0), up to a subsequence. Moreover, $(u_0, v_0) = (\hat{u}, \hat{v})$ is the solution of* (18).

Proof. First, we need to show that there is $u \in BV$ and $v \in F$ such that $\mathcal{E}_\mu(u, v) \le C$, where C does not depend on μ.

From Remark 5, the ROF model with some appropriate λ' ensures such a $u \in BV$ and a $v \in G \subset F$ such that $f = u + v$, and $\int_\Omega u = \int_\Omega f$. Therefore,

$$\mathcal{E}_\mu(u, v) = |u|_{BV} + \lambda\|v\|_F \le C,$$

and C does not depend on μ. Another pair that satisfies this property is in fact provided by $(\hat{u}, \hat{v})$.

Then we obtain that

$$\begin{gathered} \mathcal{E}_\mu(u_\mu, v_\mu) \le \mathcal{E}_\mu(u, v) \le C \\ \Rightarrow \mu\|f - u_\mu - v_\mu\|_{L^2} \le C \\ \Rightarrow \|f - u_\mu - v_\mu\|_{L^2} \le \frac{C}{\mu}; \\ \text{therefore, } \|f - u_\mu - v_\mu\|_{L^2} \to 0 \text{ as } \mu \to \infty. \end{gathered}$$

Now, as before, we can deduce that $\|u_\mu\|_{BV} \leq C$ and $\|v_\mu\|_F \leq C$. Then again, similarly, we deduce that there is (u_0, v_0) such that, up to a subsequence, $u_\mu \to u_0$ in BV-weak* and weakly in L^2, and $v_\mu \to v_0$ weakly in L^2. Moreover, and as before, we will have that

$$\begin{aligned}\mathcal{E}(u_0, v_0) &= |u_0|_{BV} + \lambda\|v_0\|_F \leq |u_\mu|_{BV} + \lambda\|v_\mu\|_F \\ &\leq \mu\|f - (u_\mu + v_\mu)\|_{L^2}^2 + |u_\mu|_{BV} + \lambda\|v_\mu\|_F \leq \mathcal{E}_\mu(\hat{u}, \hat{v}) = \mathcal{E}(\hat{u}, \hat{v});\end{aligned}$$

i.e., $(u_0, v_0) = (\hat{u}, \hat{v})$ is the minimizer of (18), and (u_0, v_0) is the limit of (u_μ, v_μ) (up to a subsequence), with $u_0 + v_0 = f$ a.e. in Ω. ∎

5.1. Characterization of minimizers. Here we would like to show some properties of minimizers of problem (19) as a generalization of Theorem 3, page 32 in [27].

We recall the variational problem of decomposing f via

$$\inf\left\{\mathcal{E}_\mu(u, v) = |u|_{BV} + \mu\int_\Omega |f - u - v|^2 \ + \lambda\|v\|_F : u \in BV, v \in F\right\}. \tag{26}$$

(Note that here we consider a larger space of possible minimizers (u, v), because we do not impose that the mean value of u is equal with the mean value of f a priori; another way would have been to work with the corresponding quotient spaces.)

DEFINITION 5.3. *Given a function $w \in L^2(\Omega)$ and $\lambda > 0$, define*

$$\|w\|_{*,\lambda} = \sup_{g\in BV(\Omega), h\in F(\Omega)\cap L^2(\Omega)} \frac{(w, g + h)}{|g|_{BV} + \lambda\|h\|_F}, \ |g|_{BV} + \lambda\|h\|_F \neq 0, \tag{27}$$

where $(\cdot, \cdot)$ is the L^2 inner product.

REMARK 6. *If $\int_\Omega w \neq 0$, then $\|w\|_{*,\lambda} = \infty$; indeed, we can replace g by $g + c$, with $c \in \mathbb{R}$, and then the supremum will no longer be finite as $|c| \to \infty$.*

We have the following characterizations of an optimal decomposition of f using (26), which will be called the (BV, F) model.

THEOREM 5.4. *Let (u, v) be an optimal (BV, F) decomposition of f, and denote $w = f - u - v$. Then we have the following:*

(1) *$\|f\|_{*,\lambda} \leq \frac{1}{2\mu} \Leftrightarrow u = 0$, $v = 0$, and $w = f$.*

(2) *Suppose $\|f\|_{*,\lambda} > \frac{1}{2\mu}$; then (u, v) is characterized by the two conditions*

$$\|w\|_{*,\lambda} = \frac{1}{2\mu} \quad \text{and} \quad (w, u + v) = \frac{1}{2\mu}(|u|_{BV} + \lambda\|v\|_F). \tag{28}$$

Proof. The (BV, F) model (26) yields $u = 0$ and $v = 0$ if and only if for any $g \in BV(\Omega)$, $h \in F(\Omega) \cap L^2(\Omega)$,

$$\mu\|f\|_{L^2}^2 \leq |g|_{BV} + \mu\|f - g - h\|_{L^2}^2 + \lambda\|h\|_F. \tag{29}$$

Equation (29) holds if and only if (by substituting in (29) g by ϵg and h by ϵh, and taking $\epsilon \to 0$)

$$|(f, g + h)| \leq \frac{1}{2\mu}(|g|_{BV} + \lambda\|h\|_F). \tag{30}$$

By the definition of $\|.\|_{*,\lambda}$, we have $\|f\|_{*,\lambda} \leq \frac{1}{2\mu}$.

For the converse property in (1), assume that $\|f\|_{*,\lambda} \leq \frac{1}{2\mu}$. Then, for any $g \in BV(\Omega)$ and $h \in F(\Omega) \cap L^2(\Omega)$, with $|g|_{BV} + \lambda\|h\|_F \neq 0$, we have

$$(f, g+h) \leq (|g|_{BV} + \lambda\|h\|_F)\|f\|_{*,\lambda} \leq \frac{1}{2\mu}(|g|_{BV} + \lambda\|h\|_F).$$

We also have

$$\begin{aligned}
&|g|_{BV} + \mu\|f - (g+h)\|_{L^2}^2 + \lambda\|h\|_F \\
&\quad = |g|_{BV} + \mu\|f\|_{L^2}^2 - 2\mu(f, g+h) + \mu\|g+h\|_{L^2}^2 + \lambda\|h\|_F \\
&\quad \geq |g|_{BV} + \mu\|f\|_{L^2}^2 - (|g|_{BV} + \lambda\|h\|_F) + \mu\|g+h\|_{L^2}^2 + \lambda\|h\|_F \\
&\quad = \mu\|f\|_{L^2}^2 + \mu\|g+h\|_{L^2}^2 \geq \mu\|f\|_{L^2}^2 = \mathcal{E}_\mu(0,0).
\end{aligned}$$

Therefore, $u = 0$ and $v = 0$ give the optimal decomposition in this case.

Now suppose $\|f\|_{*,\lambda} > \frac{1}{2\mu}$. Let (u, v) be an optimal (BV, F) decomposition. We have $u \not\equiv 0$ or $v \not\equiv 0$. For $g \in BV(\Omega)$, $h \in F(\Omega) \cap L^2(\Omega)$, and $\epsilon \in \mathbb{R}$,

(31)

$$\begin{aligned}
&|u + \epsilon g|_{BV} + \mu\|w - \epsilon(g+h)\|_{L^2}^2 + \lambda\|v + \epsilon h\|_F \geq |u|_{BV} + \mu\|w\|_{L^2}^2 + \lambda\|v\|_F \\
&\quad \Rightarrow |u|_{BV} + |\epsilon||g|_{BV} + \mu\|w - \epsilon(g+h)\|_{L^2}^2 + \lambda\left(\|v\|_F + |\epsilon|\|h\|_F\right) \\
&\qquad \geq |u|_{BV} + \mu\|w\|_{L^2}^2 + \lambda\|v\|_F \\
&\quad \Rightarrow |\epsilon||g|_{BV} + \mu\|w - \epsilon(g+h)\|_{L^2}^2 + \lambda|\epsilon|\|h\|_F \geq \mu\|w\|_{L^2}^2 \\
&\quad \Rightarrow |\epsilon||g|_{BV} + \mu\left(\|w\|_{L^2}^2 - 2\epsilon(w, g+h) + \epsilon^2\|g+h\|_{L^2}^2\right) + \lambda|\epsilon|\|h\|_F \geq \mu\|w\|_{L^2}^2.
\end{aligned}$$

Dividing both sides of the last equation by $\epsilon > 0$, we obtain

$$-2\mu(w, g+h) + \epsilon\mu\|g+h\|_{L^2}^2 + |g|_{BV} + \lambda\|h\|_F \geq 0. \tag{32}$$

Taking $\epsilon \to 0$, we obtain

$$2\mu(w, g+h) \leq |g|_{BV} + \lambda\|h\|_F \quad \forall g \in BV(\Omega),\ h \in F(\Omega) \cap L^2(\Omega).$$

Therefore,

$$\|w\|_{*,\lambda} \leq \frac{1}{2\mu}. \tag{33}$$

If we take $\epsilon \in (-1, 1)$ and replace (g, h) with (u, v) in (31), then (31) implies

$$2\mu\epsilon(w, u+v) \leq \epsilon\left(|u|_{BV} + \lambda\|v\|_F\right) + \epsilon^2\mu\|u+v\|_{L^2}^2. \tag{34}$$

If $\epsilon > 0$, $2\mu(w, u+v) \leq (|u|_{BV} + \lambda\|v\|_F)$, and if $\epsilon < 0$, $2\mu(w, u+v) \geq (|u|_{BV} + \lambda\|v\|_F)$. Therefore, equality holds that

$$(w, u+v) = \frac{1}{2\mu}\left(|u|_{BV} + \lambda\|v\|_F\right), \tag{35}$$

and (35) together with (33) implies $\|w\|_{*,\lambda} = \frac{1}{2\mu}$.

Conversely, if (35) holds for some (u, v) and $\|w\|_{*,\lambda} = \frac{1}{2\mu}$, then for any $g \in BV(\Omega)$, $h \in F(\Omega) \cap L^2(\Omega)$,

$$\begin{aligned}
&|u + \epsilon h|_{BV} + \mu\|w - \epsilon(g + h)\|_{L^2}^2 + \lambda\|v + \epsilon h\|_F \\
&\quad \geq 2\mu(w, u + \epsilon g + v + \epsilon h) + \mu\|w\|_{L^2}^2 - 2\mu\epsilon(w, g + h) + \mu\epsilon^2\|g + h\|_{L^2}^2 \\
&\quad = 2\mu(w, u + v) + \mu\|w\|_{L^2}^2 + \mu\epsilon^2\|g + h\|_{L^2}^2 \\
&\quad = |u|_{BV} + \lambda\|v\|_F + \mu\|w\|_{L^2}^2 + \mu\epsilon^2\|g + h\|_{L^2}^2 \\
&\quad \geq |u|_{BV} + \lambda\|v\|_F + \mu\|w\|_{L^2}^2.
\end{aligned}$$

Therefore, (u, v) is an optimal (BV, F) decomposition of f. ∎

REMARK 7. *Similar results also hold for the optimal (BV, G) decomposition of f, with G replacing F in* (26).

REMARK 8. *Note that if a given image f does not have zero mean, i.e., $\int_\Omega f \neq 0$, then the first property in Theorem* 5.4 *will not hold since $\|f\|_{*,\lambda} = \infty$. Therefore, if (u_μ, v_μ) is a minimizer of* (26)*, then $\|w_\mu\|_{*,\lambda} \to 0$ as $\mu \to \infty$, as expected. Here $w_\mu = f - u_\mu - v_\mu$. This is in agreement with the result from Theorem* 5.2.

5.2. Minimization of (20). For numerical computations, we use BMO^β, $BMO_\mathcal{D}$, and BMO to represent the functions g_i, $i = 1, 2$. From now on we denote BMO to mean either BMO^β, $BMO_\mathcal{D}$, or BMO, and B is either an open set, an α translated dyadic square, or a square in Ω.

Equation (20) can be further simplified as

$$\begin{aligned}
\inf\Bigg\{\mathcal{E}(u, g_1, g_2) = &\int_\Omega |\nabla u| \; + \mu\int_\Omega |f - u - \partial_x g_1 - \partial_y g_2|^2 \\
&+ \lambda\Bigg[\frac{1}{|B_1|}\int_\Omega |g_1 - g_{1,B_1}|H(\phi_1) \\
&+ \frac{1}{|B_2|}\int_\Omega |g_2 - g_{2,B_2}|H(\phi_2)\Bigg]\Bigg\},
\end{aligned} \tag{36}$$

where $g_{i,B_i} = \frac{\int_\Omega g_i H(\phi_i)}{\int_\Omega H(\phi_i)}$, ϕ_i is the level set of B_i, and B_i maximizes $\|g_i\|_{BMO}$. The infimum in (36) is taken over all $u \in BV$, and $\vec{g} = (g_1, g_2)$ with $g_i \in BMO$.

Keeping B_1 and B_2 fixed for one iteration, and minimizing $\mathcal{E}(u, g_1, g_2)$ with respect to its variables, we obtain

$$-\mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) - 2\mu(f - u - \partial_x g_1 - \partial_y g_2) = 0, \tag{37}$$

$$\begin{aligned}
&2\mu\partial_x(f - u - \partial_x g_1 - \partial_y g_2) \\
&\quad + \frac{\lambda H(\phi_1)}{|B_1|}\left[\frac{g_1 - g_{1,B_1}}{|g_1 - g_{1,B_1}|} - \frac{1}{|B_1|}\int_\Omega \frac{g_1 - g_{1,B_1}}{|g_1 - g_{1,B_1}|}H(\phi_1)\right] = 0,
\end{aligned}$$

$$\begin{aligned}
&2\mu\partial_y(f - u - \partial_x g_1 - \partial_y g_2) \\
&\quad + \frac{\lambda H(\phi_2)}{|B_2|}\left[\frac{g_2 - g_{2,B_2}}{|g_2 - g_{2,B_2}|} - \frac{1}{|B_2|}\int_\Omega \frac{g_2 - g_{2,B_2}}{|g_2 - g_{2,B_2}|}H(\phi_2)\right] = 0,
\end{aligned}$$

with the boundary conditions on $\partial\Omega$,

$$\begin{aligned}\frac{\nabla u}{|\nabla u|}\cdot\vec{n}&=0,\\ (f-u-\partial_x g_1-\partial_y g_2)n_x&=0,\\ (f-u-\partial_x g_1-\partial_y g_2)n_y&=0,\end{aligned}$$

where $\vec{n}=(n_x,n_y)$ is the exterior unit normal to $\partial\Omega$. At each iteration, for g_1 and g_2 fixed or previously estimated, the unknown sets B_i are numerically computed and updated by the methods introduced in the previous sections.

Note that by integrating both sides of (37) over Ω, the constraint in (21) is automatically satisfied.

6. Approximating the (BV,G) decomposition model. For comparison with the (BV,F) model, we approximate the (BV,G) model

$$\inf_{(u,v)\in BV\times G}\{\mathcal{E}(u,v)=|u|_{BV}+\lambda\|v\|_G\}$$

by the model

$$\begin{aligned}\inf_{(u,\vec{g})\in BV\times L^\infty(\Omega,\mathbb{R}^2)}\Bigg\{&\mathcal{E}(u,\vec{g})=|u|_{BV}+\mu\int_\Omega|f-u-\operatorname{div}(\vec{g})|^2\\ &+\lambda\int_\Omega\sqrt{g_1(x,y)^2+g_2(x,y)^2}\,\delta(x-x_0,y-y_0)\ :\\ &\int_\Omega u\ =\int_\Omega f\Bigg\}.\end{aligned}\tag{38}$$

Here $v=\operatorname{div}\vec{g}$, δ is the Dirac function (an impulse function) in two dimensions concentrated at the origin, and

$$\sqrt{g_1(x_0,y_0)^2+g_2(x_0,y_0)^2}=\||\vec{g}|\|_{L^\infty}.$$

For numerical computation, we approximate δ by a smooth version δ_ϵ such that $\delta_\epsilon\to\delta$ as $\epsilon\to 0$.

For (x_0,y_0) fixed but updated at each iteration, and minimizing $\mathcal{E}(u,\vec{g})$ with respect to u, g_1, and g_2, we obtain the Euler–Lagrange equations

$$\begin{aligned}-\operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right)-2\mu(f-u-\partial_x g_1-\partial_y g_2)&=0,\\ 2\mu\partial_x(f-u-\partial_x g_1-\partial_y g_2)+\lambda\frac{g_1}{\sqrt{g_1^2+g_2^2}}\delta_\epsilon(x_0,y_0)&=0,\\ 2\mu\partial_y(f-u-\partial_x g_1-\partial_y g_2)+\lambda\frac{g_2}{\sqrt{g_1^2+g_2^2}}\delta_\epsilon(x_0,y_0)&=0,\end{aligned}$$

with the boundary conditions

$$\begin{aligned}\frac{\nabla u}{|\nabla u|}\cdot\vec{n}&=0,\\ (f-u-\partial_x g_1-\partial_y g_2)n_x&=0,\\ (f-u-\partial_x g_1-\partial_y g_2)n_y&=0.\end{aligned}$$

For approximating the Dirac delta function, we use [13]

$$\delta_\epsilon(z)=\frac{\epsilon}{\pi(\epsilon^2+z^2)}.\tag{39}$$

7. A more isotropic (BV, F) decomposition.

REMARK 9. *$v = \operatorname{div}(\vec{g})$ is not isotropic. Also on the rectangular grid, $\operatorname{div}(\frac{\nabla u}{|\nabla u|})$ is zero at horizontal and vertical edges, and nonzero at other edges numerically. More specifically, the vertical and horizontal oscillations are not so well captured in v if we represent v as $\operatorname{div}(\vec{g})$ but are captured in the u component (as will be seen in our numerical results).*

To overcome these effects, we consider a more isotropic decomposition. As in the Osher–Solé–Vese model [33], we impose $v = \operatorname{div}(\vec{g}) = \Delta P$, for some scalar function P, to allow stronger smoothing on u. Therefore, the model (36) can be rewritten as

$$\begin{aligned}\inf\Bigg\{\mathcal{E}(u,P) = \int |\nabla u| \; &+ \mu \int_\Omega |f-u-\Delta P|^2 \\ &+ \lambda\Bigg[\frac{1}{|B_1|}\int_\Omega |P_x - P_{x,B_1}| H_\epsilon(\phi_1) \\ &+ \frac{1}{|B_2|}\int_\Omega |P_y - P_{y,B_2}| H_\epsilon(\phi_2)\Bigg]\Bigg\},\end{aligned} \tag{40}$$

where H_ϵ is a smooth approximation of the Heaviside function H, and the unknown sets B_1 and B_2 maximize the BMO norms of $g_1 = P_x$ and of $g_2 = P_y$. For fixed B_1 and B_2 but updated after each iteration, minimizing $\mathcal{E}(u,P)$ in (40) with respect to u and P, we obtain the Euler–Lagrange equations

$$\begin{aligned}&-\operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right) - 2\mu\,(f-u-\Delta P) = 0 \\ &-2\mu\Delta(f-u-\Delta P) - \frac{\lambda}{|B_1|}\left[\partial_x\left(\frac{P_x - P_{x,B_1}}{|P_x - P_{x,B_1}|}H_\epsilon(\phi_1)\right)\right] \\ &+\frac{\lambda}{|B_1|^2}\left(\int \frac{P_x - P_{x,B_1}}{|P_x - P_{x,B_1}|}H(\phi_1)\right)\partial_x H_\epsilon(\phi_1) \\ &-\frac{\lambda}{|B_2|}\left[\partial_y\left(\frac{P_y - P_{y,B_2}}{|P_y - P_{y,B_2}|}H_\epsilon(\phi_2)\right)\right] \\ &+\frac{\lambda}{|B_2|^2}\left(\int \frac{P_y - P_{y,B_2}}{|P_y - P_{y,B_2}|}H_\epsilon(\phi_2)\right)\partial_y H_\epsilon(\phi_2) = 0,\end{aligned}$$

with the boundary conditions

$$\begin{aligned}\frac{\nabla u}{|\nabla u|}\cdot\vec{n} = 0,\\ (f-u-\Delta P)n_x = 0,\ (f-u-\Delta P)n_y = 0,\\ (\nabla(f-u-\Delta P))\cdot\vec{n} = 0.\end{aligned}$$

8. Numerical results. In this section, we present numerical results for image denoising and texture decomposition obtained from the proposed models. We also show comparisons with the Vese–Osher (VO) model (4) and the ROF model (2).

Let f be the noisy version of the true image $\bar{u}$ of size $M \times N$, and let u be the denoised image. Denote

$$RMSE = \frac{\sqrt{\sum_{i=1,j=1}^{M,N}(u(i,j) - \bar{u}(i,j))^2}}{MN}.$$

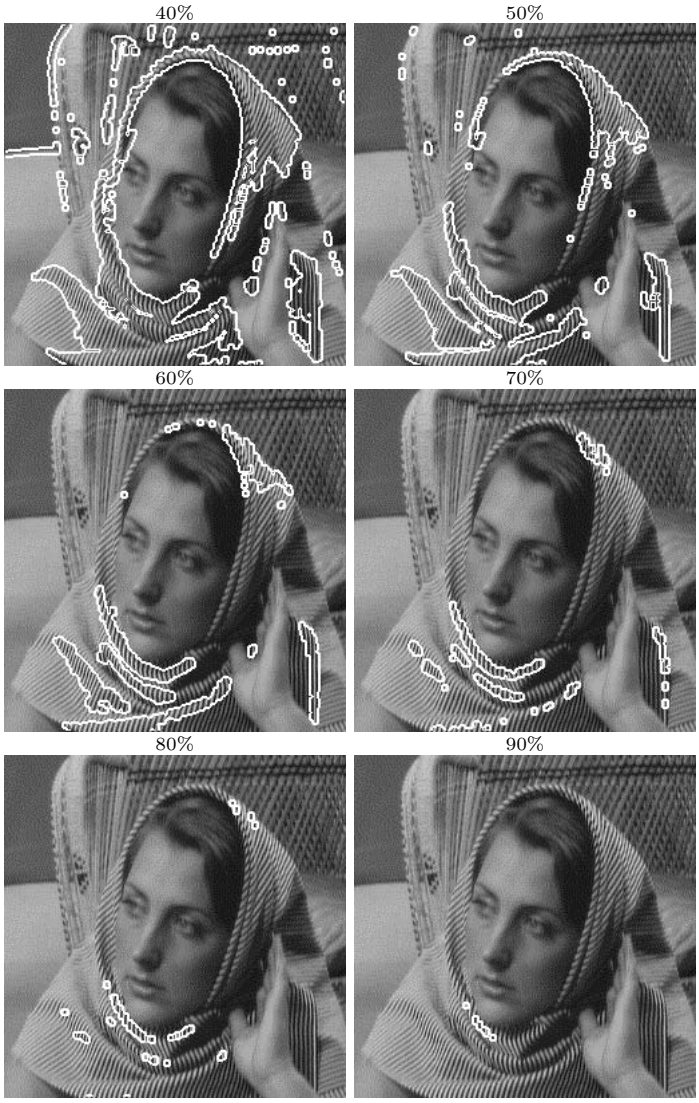


FIG. 8.1. *The union of the squares within the given percentage of the BMO norm.*

We use $RMSE$ to quantify how good a denoised image is.

Our numerical results obtained use (39) with $\epsilon = 0.01$ to approximate the Dirac delta function. We also normalize the image domain Ω, so that $\Omega = [0,1] \times [0,1]$.

In Figure 8.1, we compute the square that maximizes the BMO norm (6) and then show the contour of the union of the squares (of the same length) such that the right-hand side of (6) is within the given percentage of the BMO norm. We see that the union of the squares captures the most oscillatory regions in the given image.

Figures 8.2 and 8.3 show the contours of open sets that optimize the energies (12) and (13), respectively, using the algorithm described in section 3.1, and the plots showing the evolution of the energies versus the number of iterations.

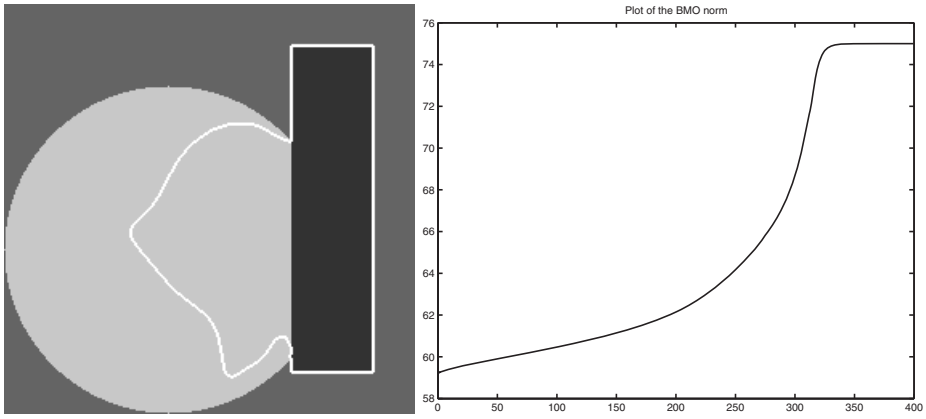


FIG. 8.2. *Left: An optimal set which gives* $\|.\|_{BMO^\beta}$. *Right: The evolution of the energy* (12) *versus iterations.*

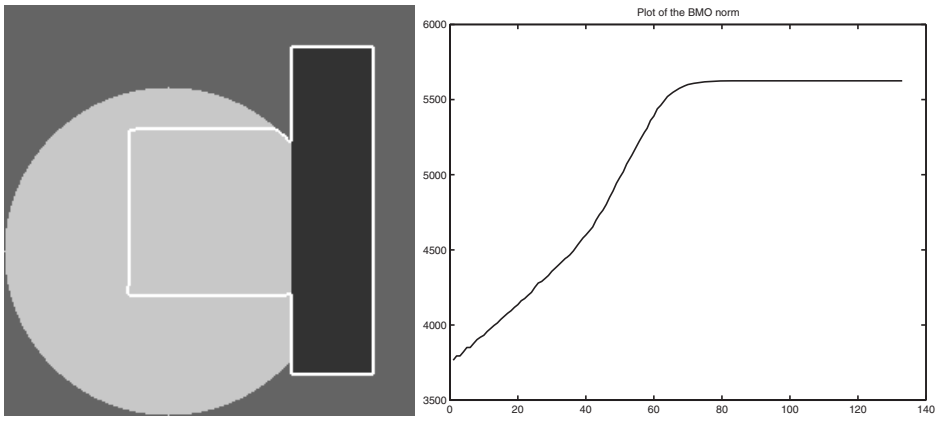


FIG. 8.3. *Left: An optimal set which gives* $\|.\|_{BMO^\beta}$. *Right: The evolution of the square of the energy* (13) *versus iterations.*

Figures 8.4 and 8.5 show the contours of squares that optimize the energies (14) and (15), respectively, using the algorithm described in section 3.2, and the plots showing the evolution of the energies versus the number of iterations.

Figure 8.6 shows the testing images that we use for our experiments.

Figure 8.7 shows two image denoisings, using the standard ROF model (2) and the VO model (4). The $RMSE$ for the ROF model is 0.00879536, and the $RMSE$ for the VO model is 0.00767165 in 2000 iterations.

Figure 8.8 shows an image denoising using the (BV, F) decomposition model (36) and the plot showing the evolution of the energy (36) with respect to the number of iterations. We use the dyadic BMO norm in this case. We also obtain similar results with the method described in section 3.3. The $RMSE$ for this decomposition is 0.0076569 in 2000 iterations.

Figure 8.9 shows an image denoising using the (BV, G) decomposition model (38) and the plot showing the evolution of the energy (38) with respect to the number of iterations. We use $\||\vec{g}|\|_{L^\infty}$ to compute the energy (38). The $RMSE$ for this decomposition is 0.0077463 in 10000 iterations. This shows that the (BV, G)

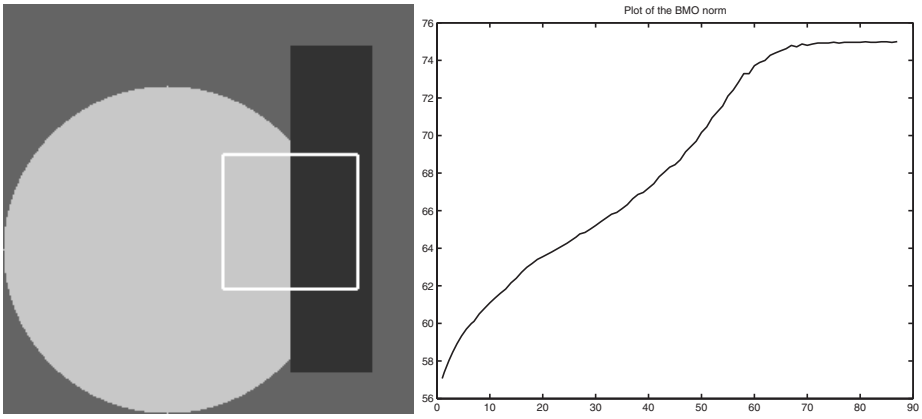


FIG. 8.4. *Left: An optimal square which gives* $\|.\|_{BMO}$. *Right: The evolution of the energy* (14) *versus iterations.*

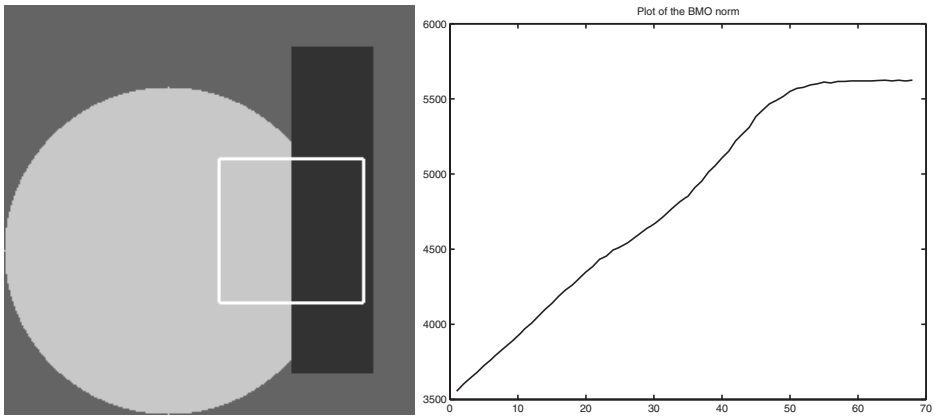


FIG. 8.5. *Left: An optimal square which gives* $\|.\|_{BMO}$. *Right: The evolution of the square of the energy* (15) *versus iterations.*

decomposition has a slower rate of convergence to the steady state in comparison with the (BV, F) decomposition. Notice that we see more of the square in the noise component $f - u$ in the ROF model than we see it in the (BV, F) and (BV, G) models.

In Figures 8.10–8.15, we show the decomposition of a given image into cartoon and texture components using (BV, F) and (BV, G) models. We remark that both models give very similar results. For the computation of the BMO norm, we use the dyadic BMO with a $\frac{1}{3}$-translations in Figure 8.10 and 8.12, and in Figure 8.14 we use the algorithm described in section 3.2 to obtain the optimal square.

Figures 8.16–8.18 show a decomposition using the standard ROF model and the proposed models. We remark that the texture parts are better captured in the oscillatory component v in Figure 8.18 using the model (40). Here we use the dyadic BMO with a $\frac{1}{3}$-translations.

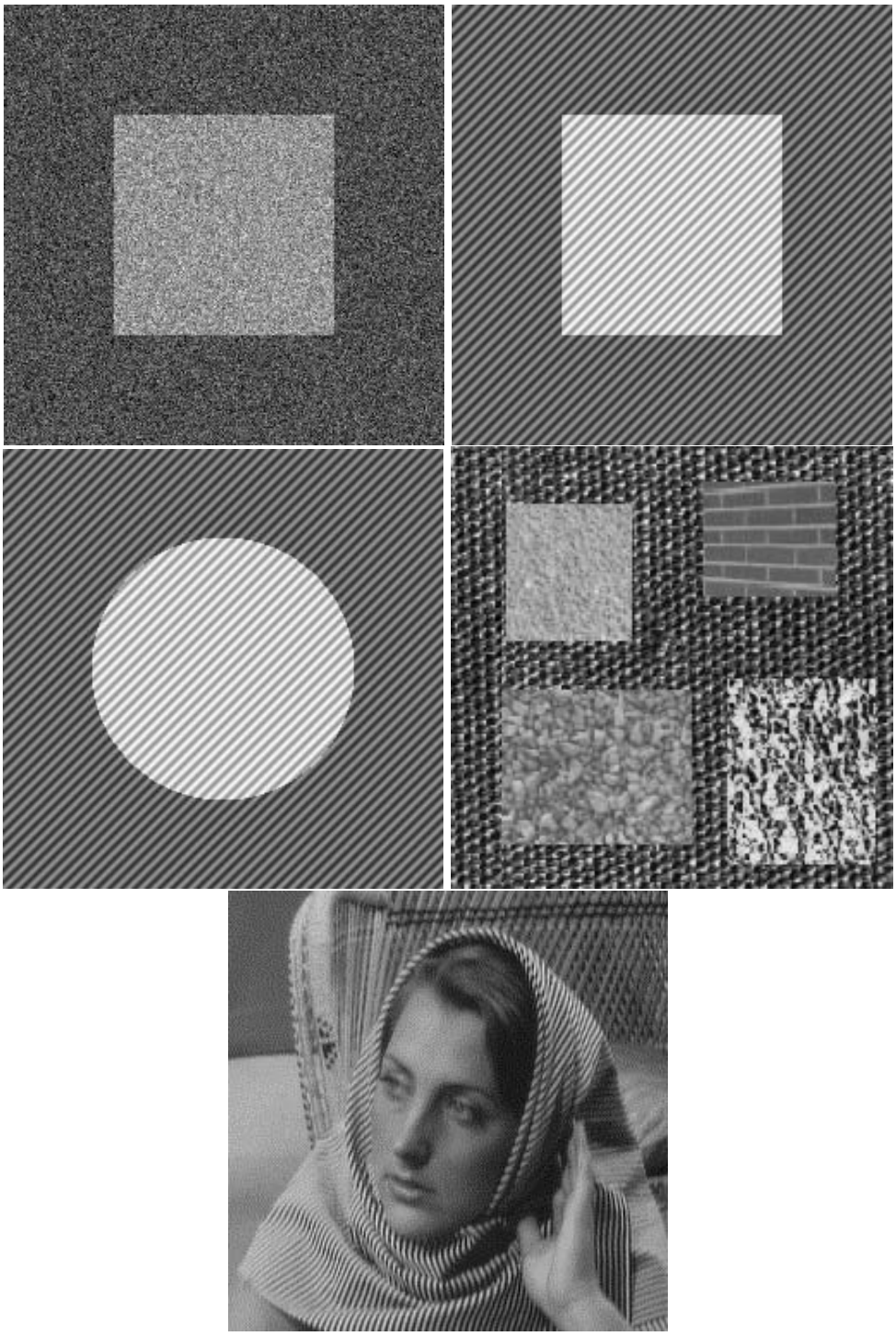


FIG. 8.6. *The data images to be decomposed.*

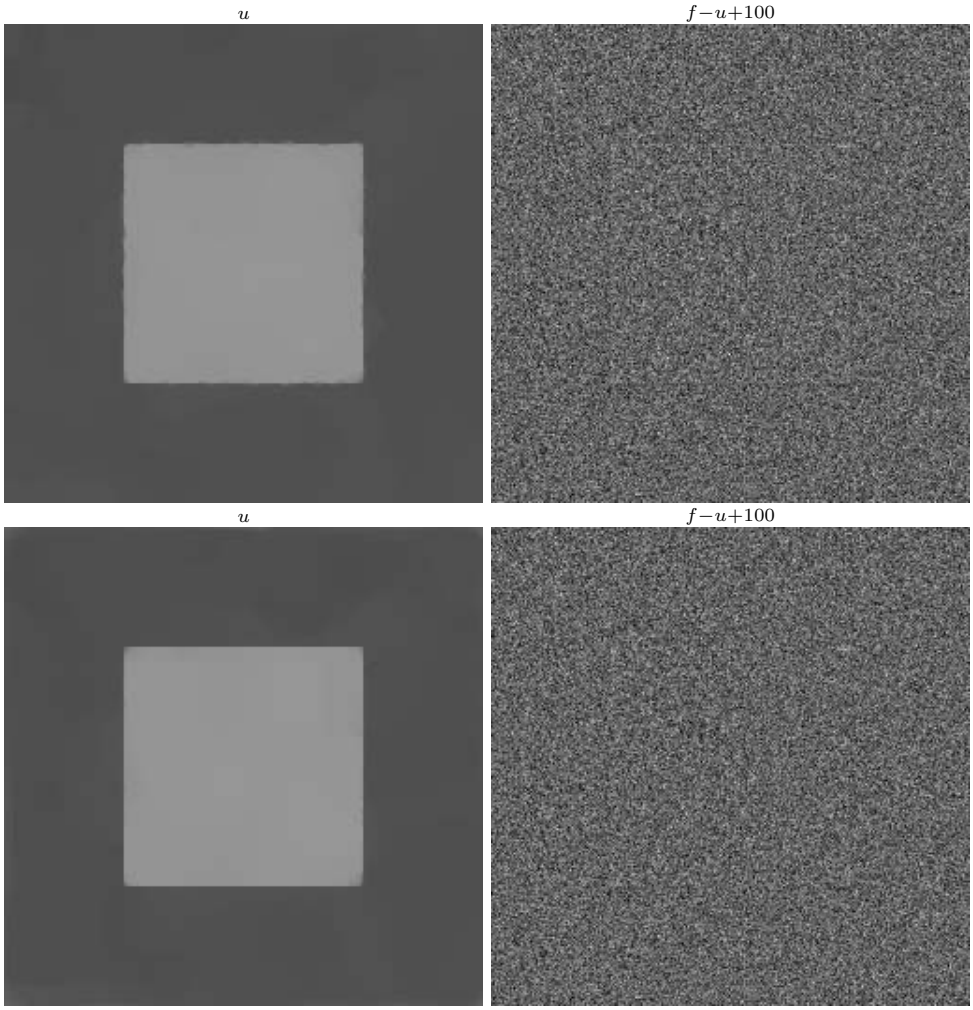
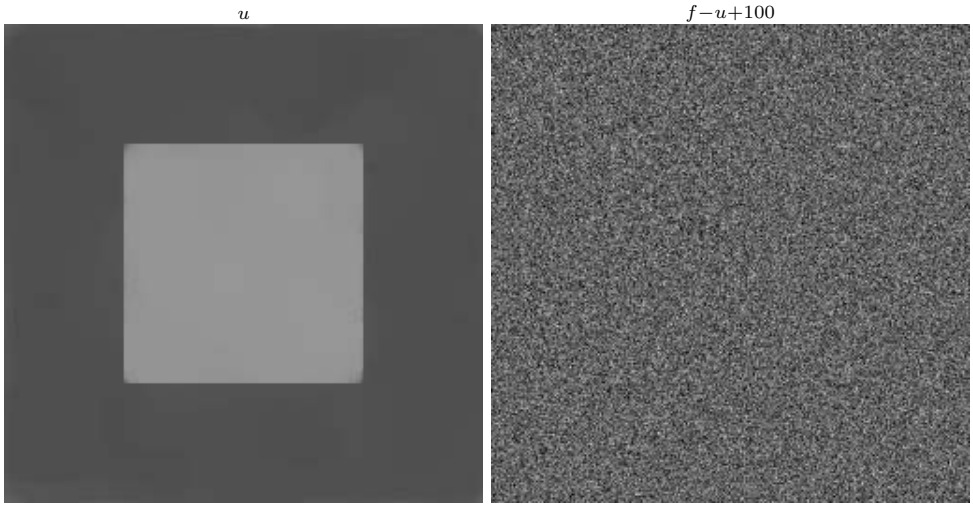


FIG. 8.7. *Top: A decomposition using the ROF model* (2). $RMSE = 0.008795368$. *Bottom: A decomposition using the VO model* (4) *with* $p = 2$, $RMSE = 0.00767165$.



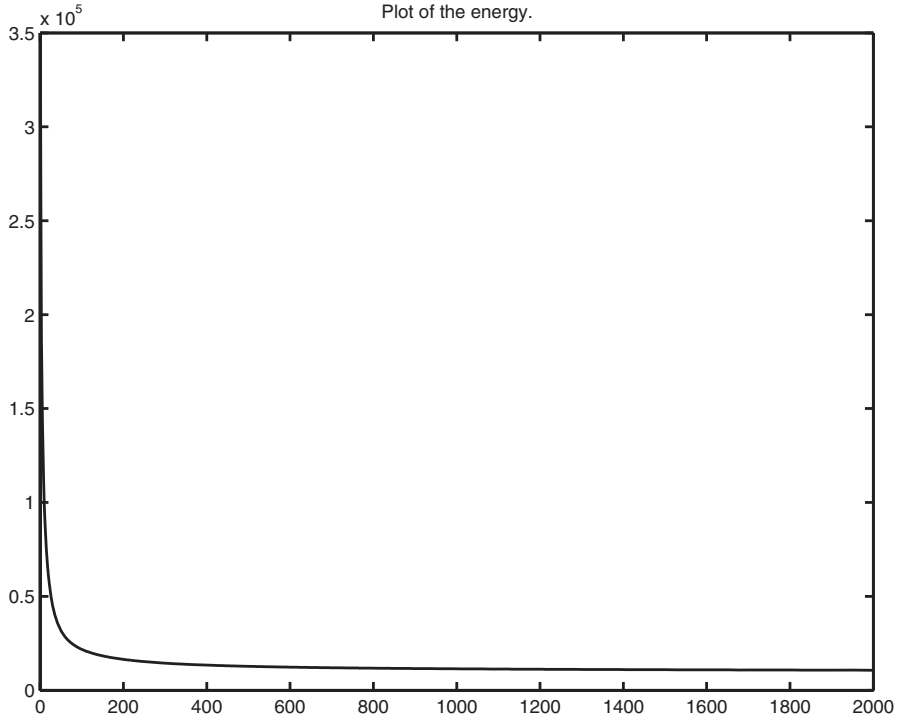


FIG. 8.8. *A decomposition using the* (BV, F) *model, with* $v = \operatorname{div}(\vec{g})$*, and the plot showing the energy* (36) *versus iterations.* $RMSE = 0.0076569$.

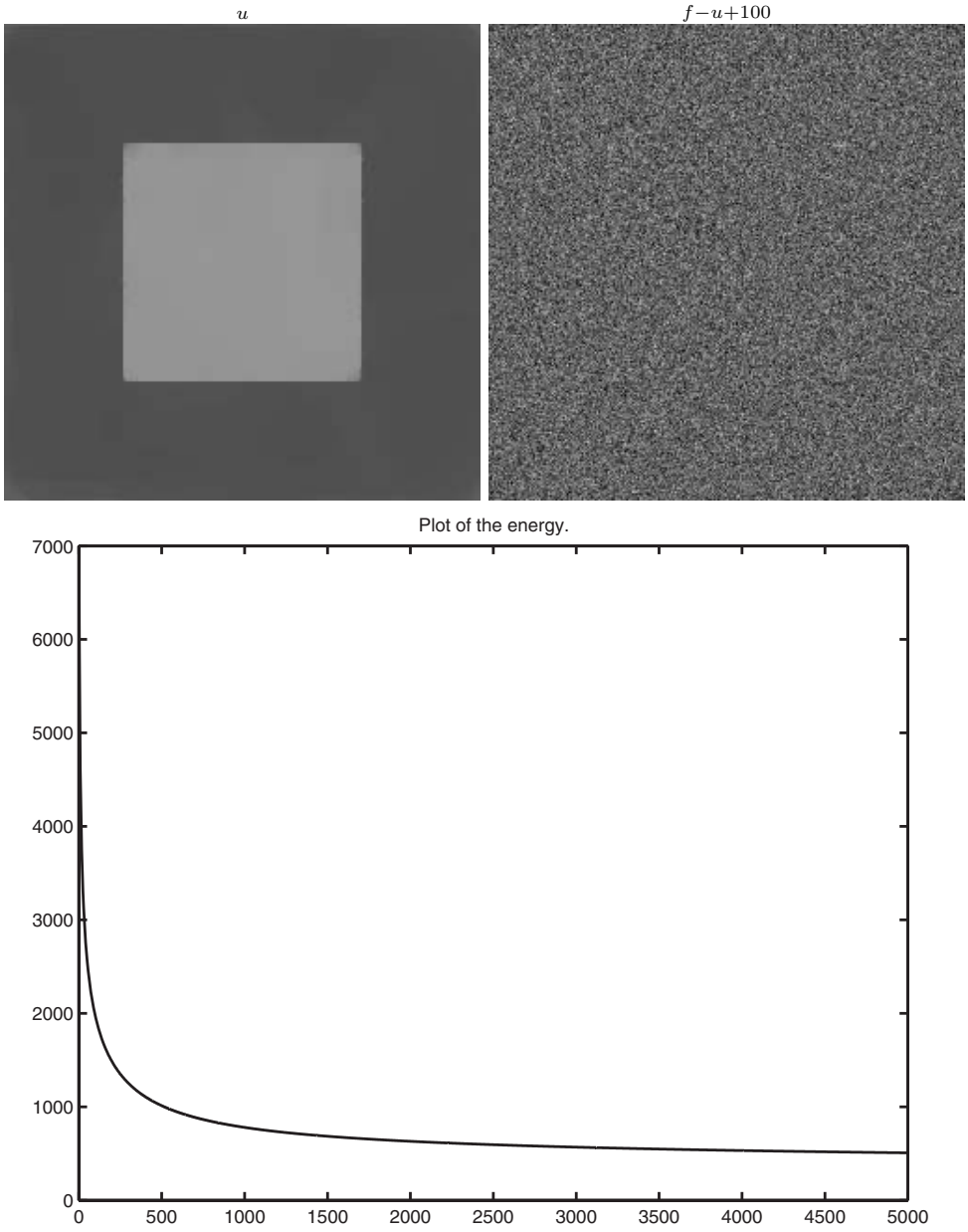


FIG. 8.9. *A decomposition using the* (BV, G) *model, with* $v = \operatorname{div}(\vec{g})$*, and the plot showing the evolution of the energy* (38) *(using* $\|\vec{g}\|_{L^\infty}$*) versus iterations.* $RMSE = 0.00775965$.

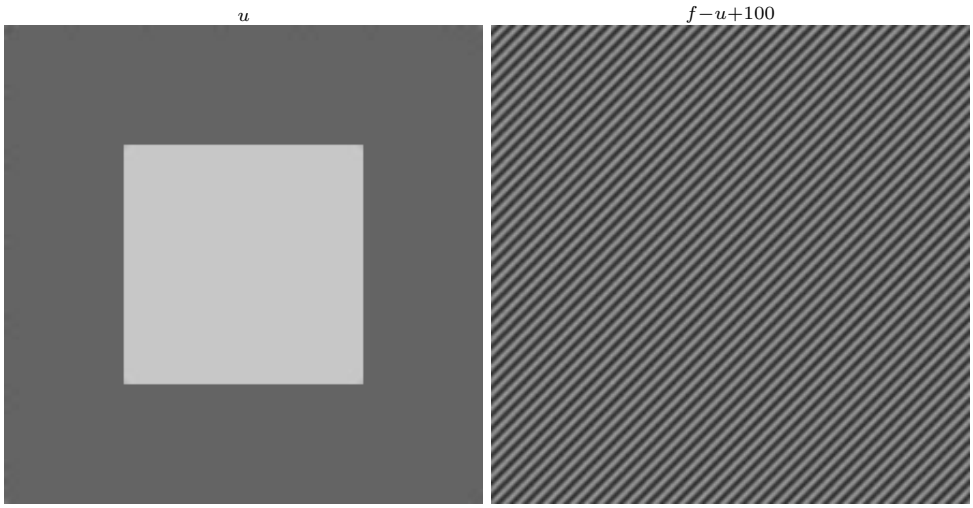


FIG. 8.10. *A decomposition using the* (BV, F) *model, with* $v = \mathrm{div}(\vec{g})$ *with translated dyadic BMO.*

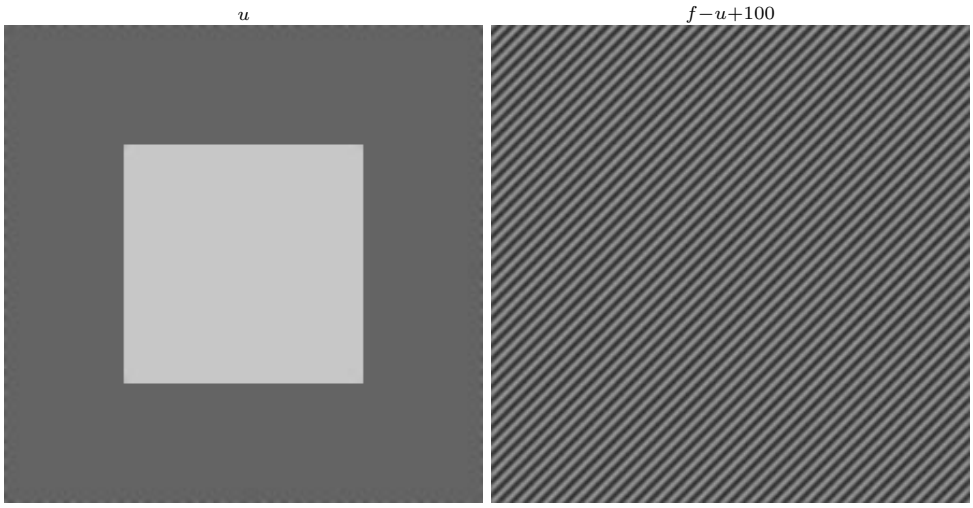


FIG. 8.11. *A decomposition using the* (BV, G) *model, with* $v = \mathrm{div}(\vec{g})$.

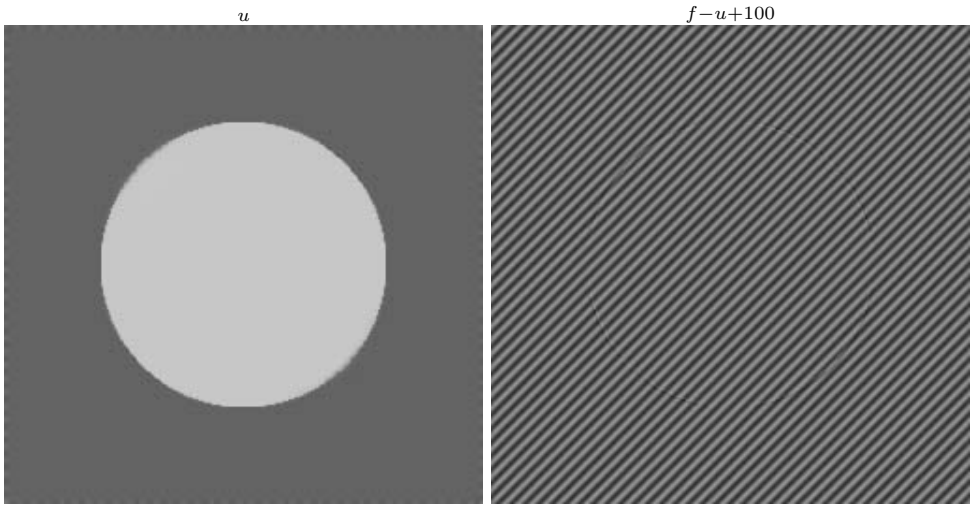


FIG. 8.12. *A decomposition using the* (BV, F) *model, with* $v = \mathrm{div}(\vec{g})$ *with translated dyadic BMO.*

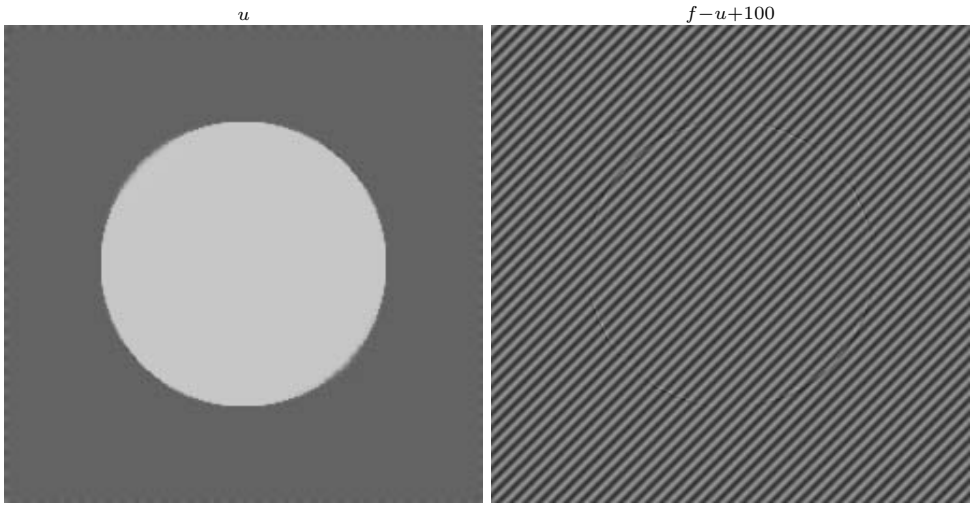


FIG. 8.13. *A decomposition using the* (BV, G) *model, with* $v = \mathrm{div}(\vec{g})$.

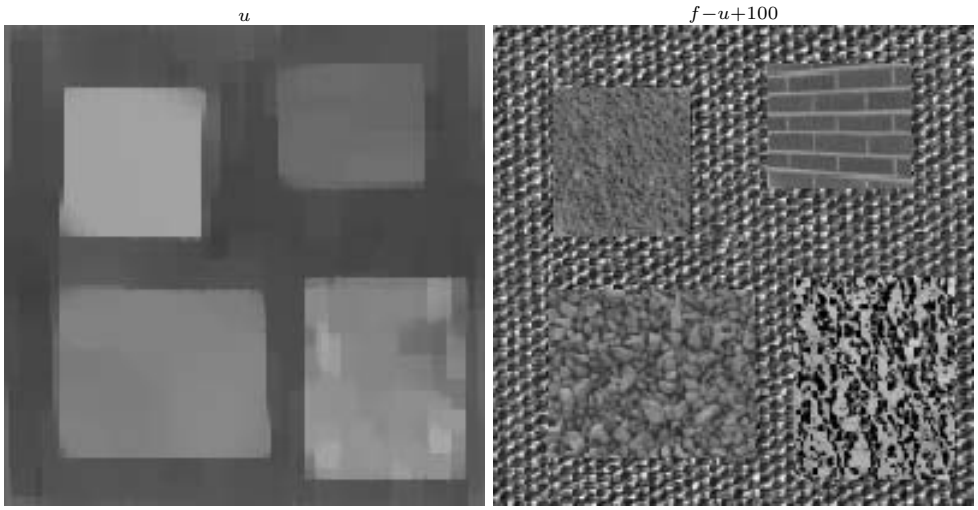


FIG. 8.14. *A decomposition using the* (BV, F) *model, with* $v = \mathrm{div}(\vec{g})$.

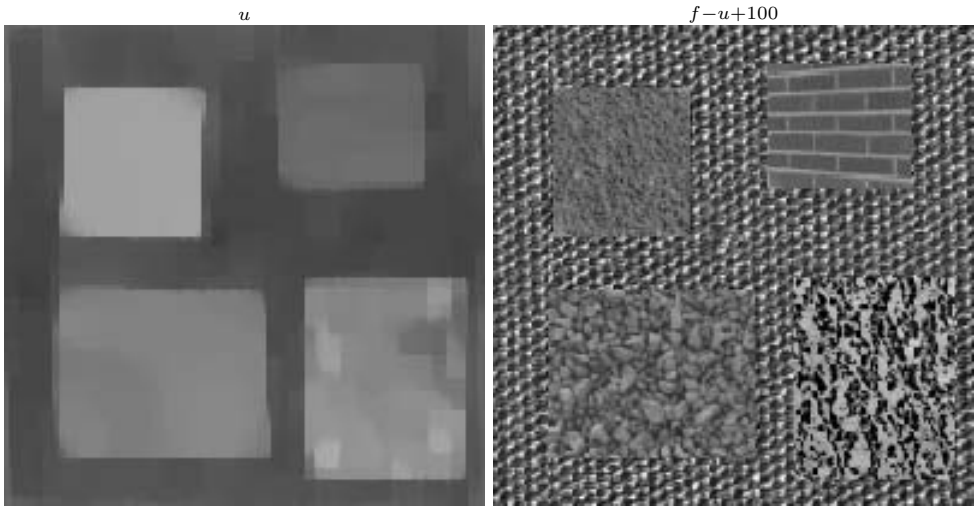


FIG. 8.15. *A decomposition using the* (BV, G) *model, with* $v = \mathrm{div}(\vec{g})$.

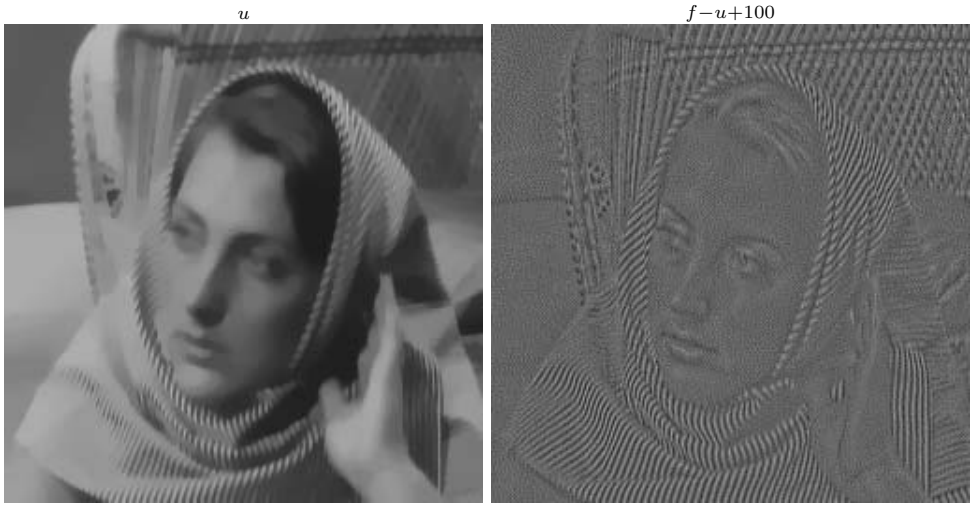


FIG. 8.16. *An ROF decomposition.*

FIG. 8.17. *A* (BV, G) *decomposition, with* $v = \mathrm{div}(\vec{g})$.

FIG. 8.18. *A decomposition using the* (BV, F) *model* (40)*, with* $v = \Delta P$.

Acknowledgments. We would like to thank John B. Garnett for the time he devoted to the authors and for his valuable and numerous remarks, discussions, and the references that he pointed out to us regarding the space BMO. We would also like to thank Peter W. Jones for his time and his valuable remarks about the space BMO and in particular for the result on dyadic $BMO_{\mathcal{D}}$ and its norm-equivalence with the usual BMO space. We would like to thank Jean-Michel Morel for his time and his valuable suggestions regarding the computation of the BMO norm and, finally, the anonymous referees for their useful suggestions that helped to improve the manuscript.

REFERENCES

[1] R. A. ADAMS, *Sobolev Spaces*, Academic Press, New York, 1975.
[2] L. ALVAREZ, Y. GOUSSEAU, AND J.-M. MOREL, *Scales in natural images and a consequence on their bounded variation norm*, in Scale-Space Theories in Computer Vision, Lecture Notes in Comput. Sci. 1682, Springer, New York, 1999, pp. 247–258.
[3] L. AMBROSIO, *Variational problems in SBV and image segmentation*, Acta Appl. Math., 17 (1989), pp. 1–40.
[4] L. AMBROSIO, N. FUSCO, AND D. PALLARA, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford University Press, New York, 2000.
[5] F. ANDREU-VAILLO, V. CASELLES, AND J. M. MAZON, *Parabolic Quasilinear Equations Minimizing Linear Growth Functionals*, Progr. Math. 223, Birkhäuser Verlag, Basel, 2004.
[6] G. AUBERT AND J.-F. AUJOL, *Modeling very oscillating signals. Application to image processing*, Appl. Math. Optim., 51 (2005), pp. 163–182.
[7] J.-F. AUJOL, G. AUBERT, L. BLANC-FÉRAUD, AND A. CHAMBOLLE, *Image decomposition application to SAR images*, in Scale-Space Methods in Computer Vision, Lecture Notes in Comput. Sci. 2695, Springer, New York, 2003, pp. 297–312.
[8] J.-F. AUJOL AND A. CHAMBOLLE, *Dual norms and image decomposition models*, Internat. J. Comput. Vision, 63 (2005), pp. 85–104.
[9] J. BOURGAIN AND H. BREZIS, *On the equation divY = f and application to control of phases*, J. Amer. Math. Soc., 16 (2002), pp. 393–426.
[10] A. CHAMBOLLE AND P.-L. LIONS, *Image recovery via total variation minimization and related problems*, Numer. Math., 76 (1997), pp. 167–188.
[11] T. F. CHAN AND S. ESEDOGLU, *Aspects of Total Variation Regularized* L^1 *Function Approximation*, UCLA CAM Report 04-07, Los Angeles, CA, 2004.

[12] T. F. Chan, S. Esedoglu, and M. Nikolova, *Algorithms for Finding Global Minimizers of Denoising and Segmentation Models*, UCLA CAM Report 04-54, Los Angeles, CA, 2004.
[13] T. F. Chan and L. Vese, *An active contour model without edges*, in Scale-Space Theories in Computer Vision, Lecture Notes in Comput. Sci. 1682, Springer, New York, 1999, pp. 141–151.
[14] T. F. Chan and L. Vese, *Active contours without edges*, IEEE Trans. Image Process., 10 (2001), pp. 266–277.
[15] E. Cheon, A. Paranjpye, L. Vese, and S. Osher, *Noise Removal Project by Total Variation Minimization*, Math 199 REU project, UCLA Department of Mathematics, Los Angeles, CA, 2002.
[16] I. Daubechies and G. Teschke, *Wavelet-based image decomposition by variational functionals*, in Wavelet Applications in Industrial Processing, Proc. SPIE 5266, F. Truchetet, ed., SPIE, Bellingham, WA, 2004, pp. 94–105.
[17] T. De Pauw, *On SBV dual*, Indiana Univ. Math. J., 47 (1998), pp. 99–121.
[18] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, CRC Press, Boca Raton, FL, 1991.
[19] S. Esedoglu and S. Osher, *Decomposition of images by the anisotropic Rudin-Osher-Fatemi model*, Comm. Pure Appl. Math., 57 (2004), pp. 1609–1626.
[20] J. B. Garnett, *Bounded Analytic Functions*, Pure Appl. Math. 96, Academic Press, New York, London, 1981.
[21] J. B. Garnett and P. W. Jones, *BMO from dyadic BMO*, Pacific J. Math., 99 (1982), pp. 351–371.
[22] D. Goldfarb and W. Yin, *Second-Order Cone Programming Methods for Total Variation-Based Image Restoration*, CORC Report TR-2004-05, Columbia University, New York, NY, 2004.
[23] Y. Gousseau and J.-M. Morel, *Are natural images of bounded variation?*, SIAM J. Math. Anal., 33 (2001), pp. 634–648.
[24] J.-L. Journe, *Calderon-Zygmund Operators, Pseudo-differential Operators and the Cauchy Integral of Calderon*, Lecture Notes in Math. 994, Springer-Verlag, Berlin, New York, 1983.
[25] H. Koch and D. Tataru, *Well-posedness for the Navier-Stokes equations*, Adv. in Math., 157 (2001), pp. 22–35.
[26] T. Mei, *BMO is the intersection of two translates of dyadic BMO*, C. R. Acad. Sci. Paris, 336 (2003), pp. 1003–1006.
[27] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, Univ. Lecture Ser. 22, AMS, Providence, RI, 2001.
[28] J.-M. Morel and S. Solimini, *Variational Methods in Image Segmentation: With Seven Image Processing Experiments*, Progr. Nonlinear Differential Equations Appl., Birkhäuser Boston, Boston, MA, 1994.
[29] D. Mumford and B. Gidas, *Stochastic models for generic images*, Quart. Appl. Math., 59 (2001), pp. 85–111.
[30] D. Mumford and J. Shah, *Optimal approximations by piecewise smooth functions and associated variational problems*, Comm. Pure Appl. Math., 42 (1989), pp. 577–685.
[31] A. Obereder, S. Osher, and O. Scherzer, *On the Use of Dual Norms in Bounded Variation Type Regularization*, UCLA CAM Report 04-35, Los Angeles, CA, 2004.
[32] S. Osher and O. Scherzer, *G-norm properties of bounded variation regularization*, Commun. Math. Sci., 2 (2004), pp. 237–254.
[33] S. Osher, A. Solé, and L. Vese, *Image decomposition and restoration using total variation minimization and the H^{-1} norm*, Multiscale Model. Simul., 1 (2003), pp. 349–370.
[34] L. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), pp. 259–268.
[35] J.-L. Starck, M. Elad, and D. L. Donoho, *Image decomposition: Separation of texture from piecewise smooth content*, in Proceedings of the SPIE Conference on Signal and Image Processing: Wavelet Applications in Signal and Image Processing X, San Diego, 2003.
[36] E. M. Stein, *Harmonic Analysis: Real Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton University Press, Princeton, NJ, 1993.
[37] R. S. Strichartz, *H^p Sobolev spaces*, Colloq. Math., 60/61 (1990), pp. 129–139.
[38] D. Strong, *Adaptive Total Variation Minimizing Image Restoration*, Ph.D. dissertation, UCLA Mathematics Department CAM Report 97-38, Los Angeles, CA, 1997.
[39] D. Strong and T. Chan, *Edge-preserving and scale-dependent properties of total variation regularization*, Inverse Problems, 19 (2003), pp. S165–S187.
[40] E. Tadmor, S. Nezzar, and L. Vese, *A multiscale image representation using hierarchical (BV, L^2) decompositions*, Multiscale Model. Simul., 2 (2004), pp. 554–579.

[41] A. Torchinsky, *Real Variable Methods in Harmonic Analysis*, Pure Appl. Math. 123, Academic Press, Orlando, FL, 1986.
[42] H. Triebel, *The Structure of Functions*, Monogr. Math. 97, Birkhäuser Verlag, Basel, 2001.
[43] L. Vese, *A study in the BV space of a denoising-deblurring variational problem*, Appl. Math. Optim., 44 (2001), pp. 131–161.
[44] L. Vese and S. Osher, *Modeling textures with total variation minimization and oscillating patterns in image processing*, J. Sci. Comput., 19 (2003), pp. 553–572.
[45] L. A. Vese and S. J. Osher, *Image denoising and decomposition with total variation minimization and oscillatory functions*, J. Math. Imaging Vision, 20 (2004), pp. 7–18.
[46] H.-K. Zhao, T. Chan, B. Merriman, and S. Osher, *A variational level set approach to multiphase motion*, J. Comput. Phys., 127 (1996), pp. 179–195.
[47] S. C. Zhu and D. Mumford, *Prior learning and Gibbs reaction-diffusion*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 19 (1997), pp. 1236–1250.