A multi-scale filament extraction method: getfilaments

Alexander Men'shchikov

A&A 560, A63 (2013)
DOI: 10.1051/0004-6361/201321885

Astronomy & Astrophysics

A multi-scale filament extraction method: *getfilaments*⋆

A. Men'shchikov

Laboratoire AIM Paris–Saclay, CEA/DSM–CNRS–Université Paris Diderot, IRFU, Service d'Astrophysique, Centre d'Études de Saclay, Orme des Merisiers, 91191 Gif-sur-Yvette, France
e-mail: alexander.menshchikov@cea.fr

Received 13 May 2013 / Accepted 6 September 2013

ABSTRACT

Far-infrared imaging surveys of Galactic star-forming regions with *Herschel* have shown that a substantial part of the cold interstellar medium appears as a fascinating web of omnipresent filamentary structures. This highly anisotropic ingredient of the interstellar material further complicates the difficult problem of the systematic detection and measurement of dense cores in the strongly variable but (relatively) isotropic backgrounds. Observational evidence that stars form in dense filaments creates severe problems for automated source extraction methods that must reliably distinguish sources not only from fluctuating backgrounds and noise, but also from the filamentary structures. A previous paper presented the multi-scale, multi-wavelength source extraction method *getsources* based on a fine spatial scale decomposition and filtering of irrelevant scales from images. Although *getsources* performed very well in benchmarks, strong unresolved filamentary structures caused difficulties for reliable source extraction. In this paper, a multi-scale, multi-wavelength filament extraction method *getfilaments* is presented that solves this problem, substantially improving the robustness of source extraction with *getsources* in filamentary backgrounds. The main difference is that the filaments extracted by *getfilaments* are now subtracted by *getsources* from detection images during source extraction, greatly reducing the chances of contaminating catalogs with spurious sources. The *getfilaments* method shares its general philosophy and approach with *getsources*, and it is an integral part of the source extraction code. The intimate physical relationship between forming stars and filaments seen in *Herschel* observations demands that accurate filament extraction methods must remove the contribution of sources and that accurate source extraction methods must be able to remove underlying filamentary structures. Source extraction with the new version of *getsources* provides researchers not only with the catalogs of sources, but also with clean images of filamentary structures, free of sources, noise, and isotropic backgrounds.

Key words. stars: formation – infrared: ISM – submillimeter: ISM – methods: data analysis – techniques: image processing – techniques: photometric

1. Introduction

In a previous paper (Men'shchikov et al. 2012, hereafter referred to as Paper I), we described the multi-scale, multi-wavelength source extraction method *getsources*. Developed primarily for large far-infrared and submillimeter surveys of star-forming regions with *Herschel*, *getsources* can also be applied to other types of astronomical images.

Instead of following the traditional approach of extracting sources directly in the observed images, *getsources* analyzes filtered single-scale decompositions of detection images over a wide range of spatial scales. The algorithm separates the peaks of real sources from those produced by the noise and background fluctuations and constructs wavelength-independent sets of combined single-scale detection images preserving spatial information from all wavebands. Sources are detected in the waveband-combined images by tracking the evolution of their segmentation masks across all scales. Source properties are measured in the observed (background-subtracted and deblended) images at each wavelength. Based on the results of an initial extraction, detection images are flattened to produce more uniform noise and background fluctuations in preparation for the second, final extraction. The method has been thoroughly tested on many simulated benchmarks and real-life images obtained in the *Herschel* Gould Belt (André et al. 2010) and HOBYS (Motte et al. 2010) surveys. The overall benchmarking results (Men'shchikov et al., in prep.) have shown that *getsources* consistently performs very well in both the completeness and reliability of source detection and the accuracy of measurements.

The wealth of high-sensitivity far-infrared images obtained with *Herschel* over the past three years have demonstrated that a substantial part of interstellar medium exists in the form of a fascinating web of omnipresent filamentary structures (see, e.g., Men'shchikov et al. 2010, for illustrations). This *anisotropic* component further complicates the very difficult problem of the systematic detection and measurement of dense cores in the strongly variable backgrounds of molecular clouds. The observational evidence that stars form in dense, cold filaments (e.g., André et al. 2010; Men'shchikov et al. 2010) creates severe problems for automated source extraction methods that must find as many *real* sources as possible from the images in several photometric bands, reliably distinguishing them not only from fluctuating backgrounds and noise, but also from the filamentary structures. The latter tend to "amplify" insignificant background or noise fluctuations that fall on top of them, confusing source extraction algorithms. The benchmarking results

⋆ Appendices are available in electronic form at http://www.aanda.org

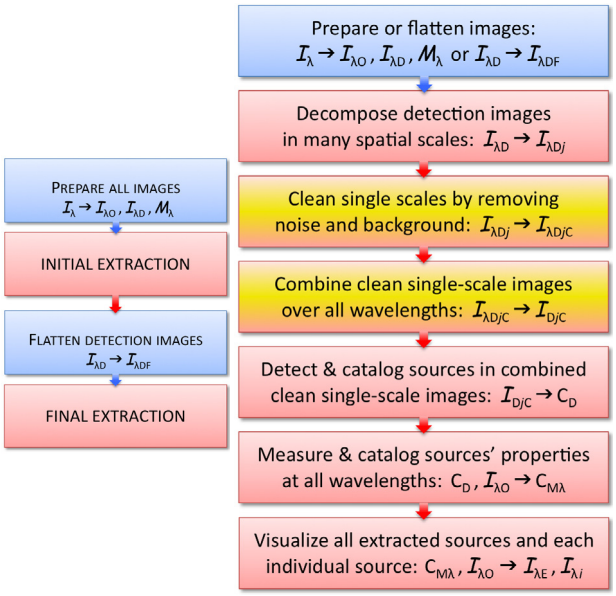

Fig. 1. Main processing blocks of *getsources* and *getfilaments*. For a *complete* extraction, both methods require two runs (*left*): the *initial* and *final* extractions (red blocks; the preparation and flattening steps are shown in blue; cf. Paper I). A more detailed presentation of the processing steps (*right*) shows that both algorithms would have had lots of identical actions, if conceived and coded separately. The *getfilaments* algorithm is essentially localized in only the cleaning and combining steps of *getsources* (highlighted in yellow). In practice, *getfilaments* is an integral part of the source extraction code, activated by a single configuration parameter of *getsources*.

(Men'shchikov et al., in prep.) suggest that source extraction methods that do not take the presence of filaments into account always tend to create significant numbers of spurious sources along the filaments.

Although *getsources* showed very good results in the benchmarks, it still created some spurious sources in simulated images with strong unresolved filamentary structures. In order to improve the performance of *getsources* in the observed *filamentary* backgrounds, a multi-scale, multi-wavelength filament extraction method *getfilaments* has been developed that solves this problem and substantially improves reliability of source extraction. The main idea behind the new approach is to carefully *extract* filaments (i.e., separate their intensity distribution from sources and largely isotropic backgrounds) and *subtract* them from the original images before detecting and measuring sources. Depending on the accuracy of the reconstructed intensity distribution, this procedure removes filamentary structures from observed images or (at least) greatly reduces their contribution. The absence of filaments in detection images makes source extraction results much more reliable, practically eliminating spurious sources.

The *getfilaments* algorithm was developed within the framework of the multi-scale and multi-wavelength approach of *getsources* (Paper I) as an integral part of the source extraction method. Both *getfilaments* and *getsources* can be described by the processing blocks shown in Fig. 1. The filament extraction method is essentially localized in only the cleaning and combining steps of *getsources*. As the *getfilaments* algorithm adds only a relatively small number of image manipulations to the original version of the source extraction method, there is no need in creating a separate code for the extraction of filaments. Moreover, the intimate physical relationship between forming stars and filaments seen in *Herschel* observations demands that accurate filament extraction methods must remove the contribution of sources and, conversely, accurate source extraction methods must be able to remove underlying filamentary structures.

This paper follows conventions and definitions introduced in Paper I. The term *noise* is used to refer to the statistical instrumental noise including possible contributions from any other signals that are not astrophysical in nature, i.e. which are not related to the emission of the areas in space one is observing. The term *background* refers to the largely isotropic astrophysical backgrounds, whereas the term *filaments* describes significantly elongated structures[1]. Filaments are *anisotropic* in the sense that their profiles and widths are very dissimilar in different directions.

Explicit distinction is made between the morphologically-simple (convex, not very elongated) *sources* of emission defined by source extraction methods and *objects* of specific astrophysical nature. In its present state, *getsources* does not know anything about the nature or true physical properties of the *objects* that produced the emission of significant peaks detected as sources. Like most of the other existing methods, it can only detect *sources* (that are possibly harboring our objects of interest) and determine their *apparent* two-dimensional intensity distributions *above* the variable background, filaments, and noise, measuring their apparent properties at each wavelength as accurately as possible.

2. Extracting filamentary structures: *getfilaments*

The fundamental problem in extracting filaments (or sources) is that all spatial scales in the images are mixed together and the intensity of any pixel contains unknown contributions from different components[2]. Following the approach formulated in Paper I, *getfilaments* analyzes decompositions of original images (in each waveband) across a wide range of spatial scales separated by a small amount (typically ~5%). Each of the "*single scales*" contains non-negligible signals from only a relatively narrow range of spatial scales, mostly only from those structures that have widths (sizes) similar to the scale considered. In effect, this automatically filters out their contributions on irrelevant (much smaller and larger) spatial scales. An immediate benefit of such filtering is that one can manipulate entire single-scale images and use thresholding to separate filaments from other structures (sources, background, and noise).

2.1. Simulated filament

The *getfilaments* algorithm is illustrated below using simulated images of a straight filament, a string of sources, a simple

[1] A quantitative definition of filaments will be formulated below (see Sect. 2.4.2), based on the areas of connected pixels occupied by structures in decomposed single-scale images. Being consistent with an intuitive idea of filaments, that formal definition results in filament lengths that are at least several times larger than their widths.

[2] There is a method that separates structural components on the basis of the wavelet, curvelet, and ridgelet decompositions (*MCA*, morphological component analysis, Starck et al. 2004). Several tests have shown that *getfilaments* gives results similar to those obtained with *MCA*. To make more detailed quantitative comparisons and conclusions, however, one has to perform an extensive benchmarking study of both methods.

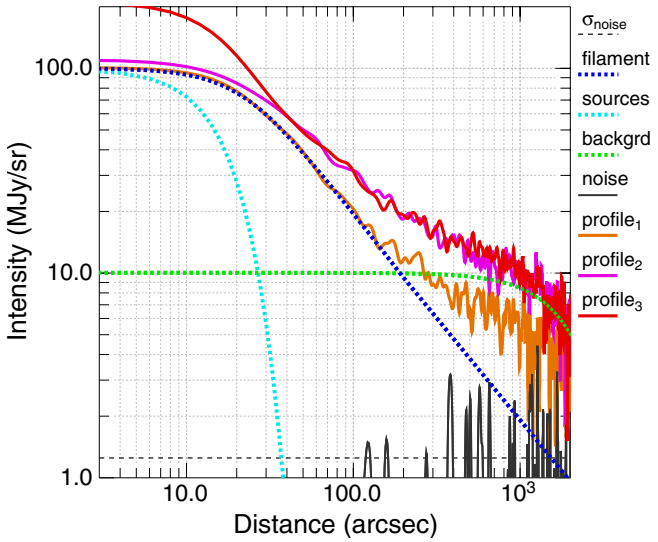

Fig. 2. Profiles of the simulated filament (Sect. 2.1) used to illustrate the *getfilaments* method. The filament has the inner half-width $R_0 = 37''$ (at half-maximum) and the outer power-law profile $I_\lambda(r) \propto r^{-1}$ (blue). Gaussian-shaped sources (cyan) and background (green) have FWHM sizes of $30''$ and $4000''$, respectively. Random pixel noise (black) was convolved to a resolution of $18''$ (FWHM) and scaled to have a standard deviation $\sigma_{\rm noise} = 1.25$ MJy/sr (dashed). Three profiles sample the full simulated image (Fig. 3) across the top of the filament (brown), just below its midpoint (magenta), and through the position of the uppermost source (red).

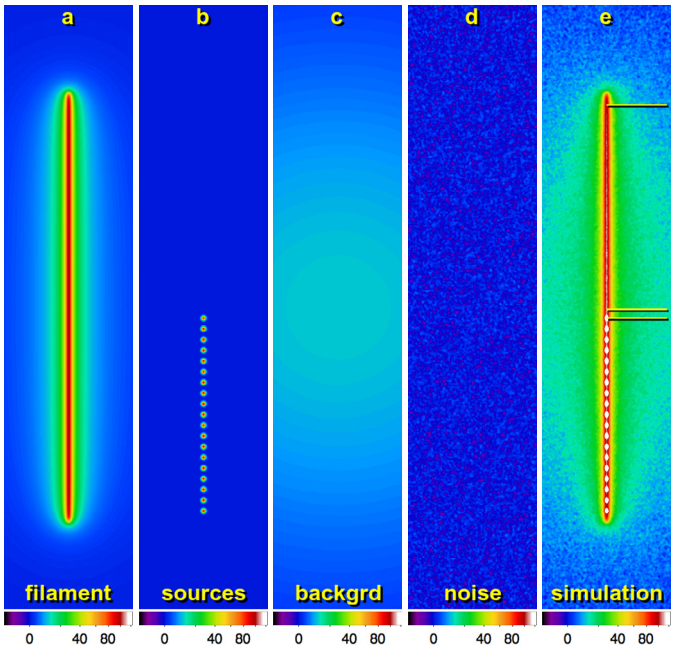

Fig. 3. Simulated image and its components (Sect. 2.1). **(a)** Straight filament with the profile displayed in Fig. 2, normalized to $\tilde{I}_{\rm P} = 100$. **(b)** Identical Gaussian sources with an intrinsic FWHM size of $24''$, convolved to $18''$ resolution and normalized to $I_{\rm P} = 100$. **(c)** Isotropic Gaussian background with a size of $4000''$ (FWHM), normalized to $I_{\rm P,b} = 10$. **(d)** Random instrumental noise with $\sigma_{\rm noise} = 1.25$. **(e)** Full simulated image $\mathcal{I}_{\lambda{\rm O}}$ ($\equiv\mathcal{I}_{\lambda{\rm D}}$) with S/N = 80 and resolution of $18''$. Three horizontal lines indicate the locations and direction of the profiles shown in Fig. 2. Each panel's dimensions are 600×2820 pixels ($0.^\circ33 \times 1.^\circ57$), pixel size $\Delta = 2''$.

background, and a moderate-level noise[3] (Figs. 2, 3), resembling the filaments observed with *Herschel* (e.g., Arzoumanian et al. 2011; Palmeirim et al. 2013). The filament profile, shown in Fig. 2, adopts the functional form (Moffat 1969) used in *getsources* to define deblending shapes (Paper I):

$$I_\lambda(r) = I_{\rm P}\left(1 + f(\zeta)\,(r/R_0)^2\right)^{-\zeta}, \qquad (1)$$

where $I_{\rm P}$ is the peak intensity, r the radial distance from the filament's crest (in the orthogonal direction), R_0 the filament's half-width at half-maximum (HWHM), ζ a power-law exponent, and $f(\zeta) = (2^{1/\zeta} - 1)$ normalizes the profile width to R_0 for all values of ζ. The function defined by Eq. (1) has Gaussian shape in its core, smoothly transforming into a power-law profile $I_\lambda(r) \propto r^{-2\,\zeta}$ for large distances $r \gg R_0$.

The parameters of Eq. (1) were fixed at $I_{\rm P} = 100$ MJy/sr, $R_0 = 18.''75$, and $\zeta = 0.5$. To simulate a $250\,\mu$m *Herschel* image, the model filament (Fig. 2) was convolved to an angular resolution of $18''$ (full width at half maximum, FWHM), preserving its peak intensity and yielding an image of the filament with a width of $D_0 = 75''$ (FWHM) and a power-law profile $I_\lambda(r) \propto r^{-1}$ at large distances (Fig. 3a). A string of identical Gaussian-shaped sources, with an intrinsic FWHM size of $24''$, were convolved to the same angular resolution of $18''$, scaled to have the same peak intensity of 100 MJy/sr, and placed along the lower half of the filament (Fig. 3b). An isotropic background was modeled as a large Gaussian ($4000''$ FWHM), normalized to 10 MJy/sr (Fig. 3c). A noise image was created by assigning random values to each pixel, convolving the resulting image to the resolution of $18''$, and scaling it to have the standard deviation $\sigma_{\rm noise} = 1.25$ MJy/sr (Fig. 3d). The simulated components were added together to produce the "observed" $250\,\mu$m image of the filament with a signal-to-noise ratio S/N = 80 (Fig. 3e). Dimensions of all images are 4800×4800 pixels ($2.^\circ66 \times 2.^\circ66$, pixel size $\Delta = 2''$), although only the central area of 600×2820 pixels, centered on the filament, is shown in this paper.

As in Paper I, images will be denoted by capital calligraphic characters (e.g., $\mathcal{A}, \mathcal{B}, \mathcal{C}$) to make a clear distinction between the images and various other parameters; all symbols and definitions are listed in Appendix A. Below, the filament extraction method (illustrated in Fig. 1) is described in full detail[4].

2.2. Preparing observed and detection images

The first step in the filament (source) extraction (Fig. 1) is to convert the original images $\mathcal{I}_\lambda$ at all wavelengths λ to the same grid and align them across wavebands, producing the observed images. This is done by resampling all images to the same (finest) pixel size (using *SWarp*, Bertin et al. 2002).

Both *getfilaments* and *getsources* distinguish between the actual observed images and detection images (denoted as $\mathcal{I}_{\lambda{\rm O}}$ and $\mathcal{I}_{\lambda{\rm D}}$, respectively). Most of the processing is done on detection images and, as the name suggests, they are used when detecting sources or filaments; observed images are used mostly for measuring and visualizing of sources. In simple cases, both $\mathcal{I}_{\lambda{\rm O}}$ and $\mathcal{I}_{\lambda{\rm D}}$ can be the same. Like *getsources* (Paper I),

[3] Simplicity of the simulation does not restrict the general applicability of *getfilaments*. Extensive experimentation has shown that the method works very well for complex, real-life filamentary fields. Numerous tests have been performed on several ground-based (sub-) millimeter images and on a dozen of multi-wavelength *Herschel* observations. See also Appendices B and C for illustrations based on very complex and realistic numerical simulations.

[4] The preparation and decomposition steps are essentially identical to those described in Paper I and, therefore, Sects. 2.2, 2.3 can safely be skipped by those familiar with *getsources*.

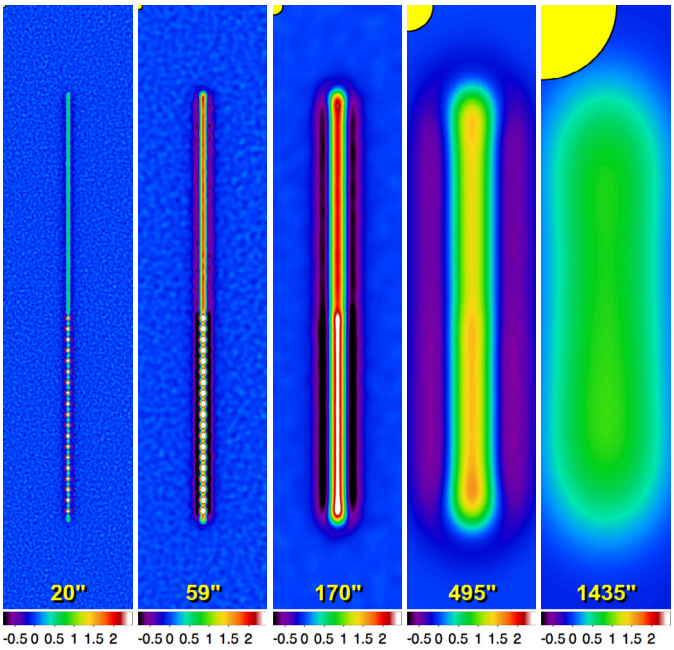

Fig. 4. Spatial decomposition (Sect. 2.3). Single scales $\mathcal{I}_{\lambda \mathrm{D}j}$ of the simulated image $\mathcal{I}_{\lambda \mathrm{D}}$ (Fig. 3) are shown for $j = 18, 32, 46, 60, 74$, $N_\mathrm{S} = 99$, $f_\mathrm{S} = 1.079$, $S_1 = 5\farcs56$, $S_{N_\mathrm{S}} = 9609''$. The scales S_j are separated by a factor of 3 to illustrate the spatial decomposition; negative areas surrounding the filament are the direct consequence of the subtraction in Eq. (2). Scale sizes S_j are visualized and annotated here and in all subsequent similar figures.

getfilaments uses convolution $\mathcal{I}_{\lambda \mathrm{D}} = \mathcal{G}_\lambda * \mathcal{I}_{\lambda \mathrm{O}}$, where $\mathcal{G}_\lambda$ is a smoothing Gaussian with its FWHM size chosen to slightly degrade (by ~5%) the image resolution O_λ. This suppresses pixel-to-pixel noise in real-life images $\mathcal{I}_{\lambda \mathrm{O}}$ (on spatial scales smaller than the observational beam size O_λ) and small-scale artifacts that would otherwise become enhanced in decomposed images.

The last part of the preparation is to create the observational masks $\mathcal{M}_\lambda$ with pixel values of either 1 or 0 that define the areas in the original images that one is interested in. They exclude from processing all pixels of $\mathcal{I}_{\lambda \mathrm{O}}$ and $\mathcal{I}_{\lambda \mathrm{D}}$ in which the mask has zero values. In the simplest case of a perfect (simulated) image, $\mathcal{M}_\lambda$ has values of 1 in all pixels. Very noisy areas (usually closer to edges) can affect the cleaning and detection algorithms and one needs to exclude them using carefully-prepared observational masks. The mask images $\mathcal{M}_\lambda$ should not have isolated holes: all zero pixels must be connected to each other and to the image edges.

2.3. Decomposing detection images in spatial scales

The spatial decomposition is done by convolving the original images with circular Gaussians of progressively larger sizes and subtracting them from one another (Fig. 4):

$$\mathcal{I}_{\lambda \mathrm{D}j} = \mathcal{G}_{j-1} * \mathcal{I}_{\lambda \mathrm{D}} - \mathcal{G}_j * \mathcal{I}_{\lambda \mathrm{D}} \ (j = 1, 2, \ldots, N_\mathrm{S}), \tag{2}$$

where $\mathcal{I}_{\lambda \mathrm{D}}$ is the detection image (Sect. 2.2), $\mathcal{I}_{\lambda \mathrm{D}j}$ are its "single-scale" decompositions, and $\mathcal{G}_j$ are the smoothing Gaussian beams ($\mathcal{G}_0$ is a two-dimensional delta-function). The beams have FWHM sizes $S_j = f_\mathrm{S}\, S_{j-1}$ in the range $2\,\Delta \lesssim S_j \lesssim S_{\max}$, where Δ is the pixel size, $f_\mathrm{S} > 1$ is the scale factor, and $S_{\max}$ is the maximum spatial scale considered. The number of scales N_S depends on the values of f_S (typically ≈ 1.05) and $S_{\max}$. The value of $S_{\max}$ is determined by the maximum sizes of filaments (sources) in the extraction and its upper limit is the size of the image along its smallest dimension. For large values of f_S, the single scales actually contain mixtures of wide ranges of scales, where faint small-scale structures become completely diluted by the contribution of irrelevant (much larger) scales. Smaller values of f_S ensure better spatial resolution of the set of single scales, just like fine mesh sizes better resolve structures in numerical methods. For values f_S that are too close to unity, images on scales j and $j + 1$ become almost identical[5].

Equation (2) implicitly assumes that the convolved images are properly rescaled to conserve their total flux; therefore, the original image can be recovered by summing up all scales:

$$\mathcal{I}_{\lambda \mathrm{D}} = \sum_{j=1}^{N_\mathrm{S}} \mathcal{I}_{\lambda \mathrm{D}j} + \mathcal{G}_{N_\mathrm{S}} * \mathcal{I}_{\lambda \mathrm{D}}. \tag{3}$$

Before convolution, the images $\mathcal{I}_{\lambda \mathrm{D}}$ are expanded from the edges of the areas covered by the observational masks $\mathcal{M}_\lambda$ towards the image edges and the entire images are expanded on all sides by a large enough number of pixels ($2\,S_j/\Delta$) to avoid undesirable border effects. Both expansions are performed using the pixel values at the edges of the masks and images, respectively, and extrapolating them outwards in four main directions (horizontal, vertical, and two diagonals). After convolution, the images are reduced back to their original size.

The single-scale decomposition of Eq. (2) filters out emission on all irrelevant scales and thus $\mathcal{I}_{\lambda \mathrm{D}j}$ reveal structures with a much higher contrast than $\mathcal{I}_{\lambda \mathrm{D}}$ does. The decomposition naturally selects filaments (sources) of specific widths, which become most visible in the images containing similar scales. The negative areas surrounding bright filaments in Fig. 4 are the direct consequence of the successive unsharp masking in Eq. (2), i.e., the subtraction of an image convolved with a larger smoothing beam from the one convolved with a smaller beam.

2.4. Cleaning single scales of noise, background, and sources

Before one can use the single-scale detection images $\mathcal{I}_{\lambda \mathrm{D}j}$ for filament extraction, they must be cleaned of the contributions of noise, background, and sources to make sure that most (if not all) non-zero pixels belong to real filamentary structures.

2.4.1. Iterative cleaning algorithm

As in Paper I, single-scale cleaning is done by the global intensity thresholding of $\mathcal{I}_{\lambda \mathrm{D}j}$. Unlike the original images $\mathcal{I}_{\lambda \mathrm{O}}$ or $\mathcal{I}_{\lambda \mathrm{D}}$ that often have a very strong and highly variable background, the entire single-scale images are "flat" in the sense that signals on considerably larger scales have been removed or greatly suppressed (see Fig. 4). Another advantage of this *single-scale cleaning* is that the noise contribution depends very significantly on the scale. For example, the small-scale noise gets heavily diluted on large scales, where extended sources become most visible. In effect, in the reconstructed clean images $\mathcal{I}_{\lambda \mathrm{D}\,\mathrm{C}} = \sum_j \mathcal{I}_{\lambda \mathrm{D}j\mathrm{C}}$, one can see large structures better (deeper) than in $\mathcal{I}_{\lambda \mathrm{D}}$.

Paper I described an iterative cleaning algorithm that automatically finds (on each scale) a cut-off level separating significant signals from those of the noise and background. On the first scale ($j = 1$), it computes the cut-off (threshold) $\varpi_{\lambda j} = n_{\lambda j}\, \sigma_{\lambda j}$, where $\sigma_{\lambda j}$ is the standard deviation over the entire image

[5] In the current implementation of the method, the minimum value of f_S is set to 1.03, whereas the maximum value of N_S is 99.

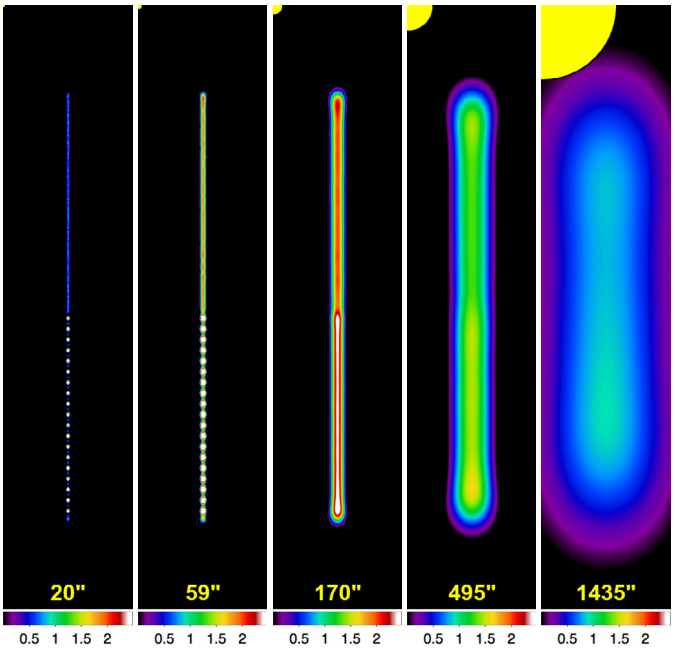

Fig. 5. Single-scale removal of noise and background (Sect. 2.4). The same set of spatial scales (Fig. 4) is displayed in the single-scale images $\mathcal{I}_{\lambda \mathrm{D}j\mathrm{C}}$, cleaned of noise and background with an iterative procedure described in full detail in Paper I. Noise fluctuations visible in Fig. 4 have been removed by zeroing pixels with $I_{\lambda j} < \varpi_{\lambda j}$.

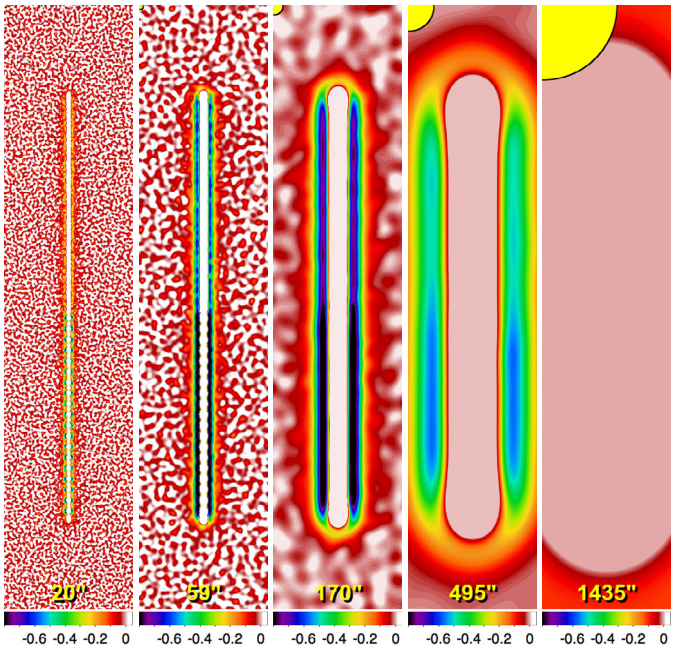

Fig. 6. Single-scale detection of filaments (Sect. 2.4). The images of spatial scales from Fig. 4 are shown here below the intensity threshold $\tilde{\varpi}_{\lambda j} = \sigma_{\lambda j}$ for detection of filaments. The filament's base is clearly visible in the single scales, although the latter are strongly contaminated by noise and background fluctuations.

$\mathcal{I}_{\lambda \mathrm{D}j}\,\mathcal{M}_{\lambda}$, and $n_{\lambda j}$ is a variable factor having an initial value of $n_{\lambda 1} = 6$. Then the procedure masks out all pixels with the values $|I_{\lambda j}| \geq \varpi_{\lambda j}$ and repeats the calculation of $\sigma_{\lambda j}$ over the remaining pixels, estimating a new threshold, which is generally lower than the one in the previous iteration. The procedure masks out bright pixels again and iterates further, always computing $\sigma_{\lambda j}$ at $|I_{\lambda j}| < \varpi_{\lambda j}$, *outside* the peaks and hollows, until $\varpi_{\lambda j}$ converges ($\delta\varpi_{\lambda j} < 1\%$) to a stable threshold (see Paper I).

Having obtained the single-scale thresholds $\varpi_{\lambda j}$ distinguishing between the significant and insignificant signals, one can create *clean* single-scale images $\mathcal{I}_{\lambda \mathrm{D}j\mathrm{C}}$, where all faint pixels with $I_{\lambda j} < \varpi_{\lambda j}$ are zeroed. This (ideally) leaves non-zero only those pixels that belong to significantly bright structures (sources, filaments). The resulting clean single-scale images of the simulated filament are illustrated in Fig. 5.

When *sources* are being extracted, $\varpi_{\lambda j}$ is the deepest level that *getsources* can descend to ($2.5\,\sigma_{\lambda j} \leq \varpi_{\lambda j} \leq 6\,\sigma_{\lambda j}$, cf. Paper I). At fainter levels, there is no reliable way of distinguishing between sources and peaks produced by noise and background fluctuations, while there is a real danger of creating spurious sources. In the case of *filaments* considered in this paper, it is possible to use the fact that filaments are substantially elongated structures (as opposed to *sources* that are not very elongated, Sect. 1) and analyze much fainter signals. Numerous tests have shown that $\tilde{\varpi}_{\lambda j} = \sigma_{\lambda j}$ is a good choice for detecting filaments. Very faint filaments are only found in the intermediate range of intensities ($\tilde{\varpi}_{\lambda j}$ to $\varpi_{\lambda j}$), where they are strongly eroded by noise and background fluctuations, as well as altered by sources. Considerably brighter filaments that rise well above the source detection threshold $\varpi_{\lambda j}$ (Fig. 5) are much less affected by fluctuations at the $\sigma_{\lambda j}$ level.

2.4.2. Cleaning algorithm for detecting filaments

Paper I has emphasized great benefits of detecting sources in single-scale images: the spatial decomposition of Eq. (2) is based on convolution and the latter acts as a natural selector of scales in decomposed images (cf. Sect. 2.3). As a consequence, resolved isolated circular sources with a FWHM size A would have their maximum peak intensity in single-scale images with smoothing beams $S_j \approx A$. Indeed, convolving with small beams ($S_j \ll A$) would have almost no effect on the source, whereas using extended beams ($S_j \gg A$) would greatly dilute the source. At both these extremes, spatial decomposition produces decreasing peak intensities, while creating the strongest signal for the sources with sizes $A \approx S_j$. Completely unresolved sources are the brightest on spatial scales $S_j \lesssim O_{\lambda}$. In effect, sizes of all significant structures seen in single-scale images are very similar to the size S_j of the smoothing beam.

Lengths L of filaments are significantly greater (at least several times) than their widths W, which makes their single-scale properties quite different from those of sources, allowing one to distinguish them from the contributions of all other components (noise, background, and sources). The spatial decomposition of Eq. (2) selects the filaments with *widths* similar to the smoothing beam ($W \approx S_j$), whereas their greater *lengths* ($L \gg S_j$) are mostly unaffected by the convolution. This means that filaments occupy much larger areas in single-scale images than any contribution from sources and fluctuations of noise or largely isotropic backgrounds.

When clipped at the filament detection threshold $\tilde{\varpi}_{\lambda j} = \sigma_{\lambda j}$, the single-scale images with intensities $I_{\lambda j} \leq \tilde{\varpi}_{\lambda j}$ clearly display the base of filamentary structures (Fig. 6). The structures can also be seen in a much simpler way (as mask images) when one zeroes all pixels with intensities $I_{\lambda j} < \tilde{\varpi}_{\lambda j}$ (Fig. 7). The low level of thresholding leads to strong contamination by noise peaks (in general, also by background fluctuations and sources) that have to be removed before the images could be used for filament extraction. Such cleaning is a simple procedure based on the comparison of the area of connected non-zero pixels with the area of the smoothing beam.

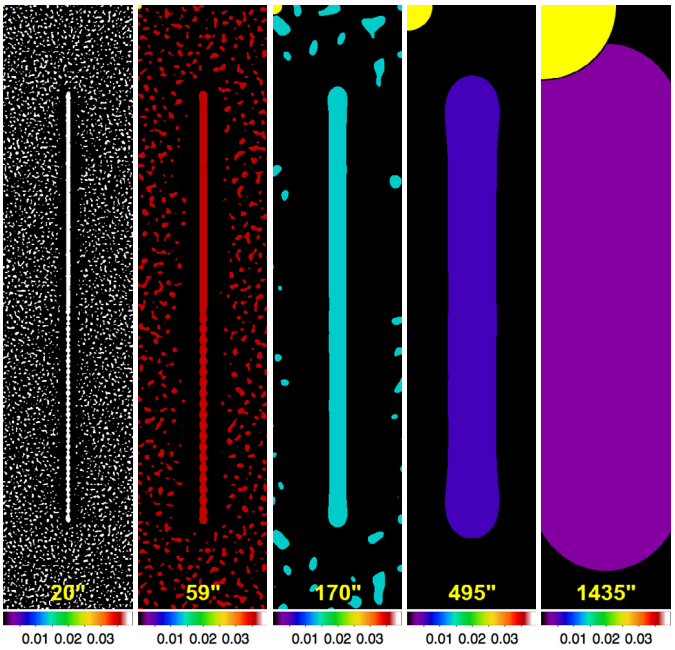

Fig. 7. Single-scale masking of filaments (Sect. 2.4). The images of spatial scales from Fig. 4 are shown here exactly on the intensity threshold $\tilde{\varpi}_{\lambda j} = \sigma_{\lambda j}$ for detection of filaments, and all lower level signals were set to zero. The simulated filament is clearly visible on all scales, as is the abundant contamination of the images by noise. Noise-free black zones that appear to surround the filament are the consequence of the negative areas seen in Figs. 4 and 6.

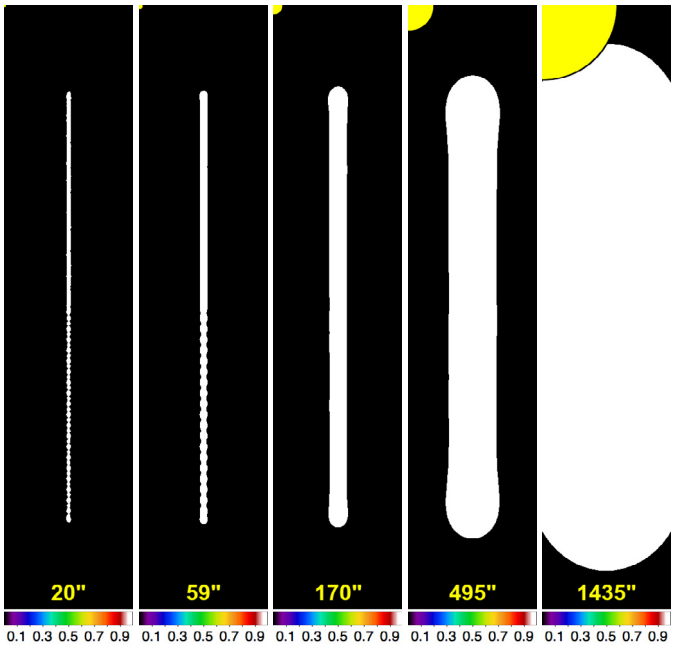

Fig. 8. Single-scale masks of filaments (Sect. 2.4). The images of spatial scales from Fig. 7 are shown here at their base level as the normalized filament masks $\mathcal{M}_{\lambda j}$ after the removal of relatively small clusters of pixels of non-filamentary nature (including noise, background, and sources) with $N_{\Pi\lambda j} < N^{\min}_{\Pi\lambda j}$. In the five images displayed, structures with less than 1.31×10^3, 8.15×10^4, 6.85×10^5, 5.76×10^6, and 5.65×10^7 pixels have been removed (cf. Fig. 7).

As illustrated in Fig. 7, non-filamentary (insignificantly elongated) structures always occupy relatively small areas, when decomposed into single-scale images and considered above the threshold level $\tilde{\varpi}_{\lambda j}$. The decomposition of Eq. (2) naturally selects structures with characteristic scales similar to S_j, filtering out both much smaller and much larger scales. Above the level $\tilde{\varpi}_{\lambda j}$, sufficiently bright filaments connect relatively large areas of pixels, because it is the filament *width* that becomes similar to the decomposition scale S_j. The longer dimension of filaments is practically unaffected by the convolution at S_j and thus stays almost the same over a much wider range of scales (cf. Fig. 4).

Cleaning of the single-scale images of noise, background, and sources is done with the *TintFill* algorithm (Smith 1979)[6] used in *getsources* for detecting sources (Paper I). The algorithm finds clusters of pixels connected to each other by their sides and fills all the pixels with a new value[7]. To remove connected clusters from images, the new value is set to zero.

Distinguishing between the real filamentary structures and smaller peaks of non-filamentary nature, in order to remove the latter from single-scale images, *getfilaments* employs a lower limit $N^{\min}_{\Pi\lambda j}$ on the number of connected pixels $N_{\Pi\lambda j}$ that are allowed to remain in the clean images of filaments:

$$N^{\min}_{\Pi\lambda j} = \tilde{f}\, N_{\rm B}\, \pi \left(3O^6_\lambda + S^6_j\right)^{2/6} \Delta^{-2}, \tag{4}$$

where O_λ is the observational beam size, S_j the smoothing (decomposition) beam, Δ the pixel size, $N_{\rm B}$ the number of cleaning beam areas ($N_{\rm B} = 30$), and $\tilde{f}$ a shape factor defined in Eq. (5) below (assume $\tilde{f} = 1$ for a moment). Clusters of pixels with $N_{\Pi\lambda j} < N^{\min}_{\Pi\lambda j}$ are removed from the decomposed images on each spatial scale (see Fig. 8).

The combination of two beams in Eq. (4) defines the effective cleaning beam designed to change smoothly and rapidly between the regimes of small and large spatial scales. On small scales, the cleaning beam becomes almost constant (approaching $1.2O_\lambda$), which is necessary to offset enhanced noisiness of small-scale images and minimize the chances of false detections. This raises the effective beam substantially above S_j on small scales ($S_j \lesssim O_\lambda$), which may lead to removal of small real filaments. To recognize filaments in small-scale structures better, one can examine shapes of the latter in addition to their areas.

The shape factor $\tilde{f}$ in Eq. (4) is designed to fine-tune $N^{\min}_{\Pi\lambda j}$ depending on various shapes of structures. To quantify them, *getfilaments* employs images of masks (Fig. 7), defining an ellipse for each cluster of connected pixels by computing their major and minor sizes (a, b) from intensity moments (e.g., Appendix F in Paper I). Simple, relatively straight filaments can be quantified by their elongation $\tilde{E}$, which is defined as the ratio a/b. However, most of the actual filaments observed with *Herschel* are curved, warped, twisted, or shaped irregularly otherwise, reflecting complex dynamical (possibly violent) processes that created them. Elongation $\tilde{E}$ alone cannot be used to quantify strongly curved, not very "dense" clusters of connected pixels that meander around (e.g., a spiral structure). To describe such a filament, one can define *sparsity* $\tilde{S}$ as the ratio of the elliptical area πab to the total area occupied by all non-zero pixels belonging to the filament. Although $\tilde{E}$ may well be very close to unity for sparse clusters of connected pixels, a high value of $\tilde{S}$ for such structures would indicate that they are filaments.

The above considerations, along with some experimentation, led to the following empirical definition of the shape factor:

$$\tilde{f} = \left(\tilde{E} \max\left\{\tilde{S}, 1\right\}\right)^{-1} \exp\left(20\,(1.2 - \tilde{E}) + 1\right). \tag{5}$$

[6] Available at http://portal.acm.org/citation.cfm?id=800249.807456

[7] Identification of distinct connected regions in similar algorithms is also known as *connected-component labeling*.

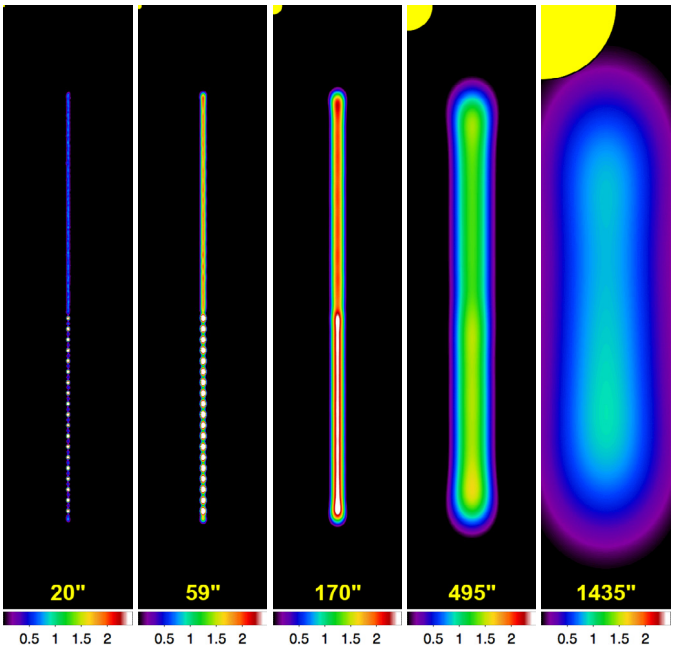

Fig. 9. Single-scale intensities within filament masks (Sect. 2.4). The images of spatial scales from Fig. 4 are shown here in pixels with $I_{\lambda j} > \tilde{\varpi}_{\lambda j}$. Roughly representing the intensity distribution of clean filaments, such images do not take negative areas into account, and they are contaminated by the emission of bright sources.

The elongation $\tilde{E}$ and sparsity $\tilde{S}$ lower the value of $\tilde{f}$, hence the required number of connected pixels in Eq. (4) for structures that are increasingly elongated and sparse. Besides, the exponential factor in Eq. (5) raises a steep barrier for structures with $\tilde{E} \lesssim 1.3$.

With the shape factor defined in Eq. (5), the simple area condition of Eq. (4) works very well for all simulated and *Herschel* images tested. The cleaning procedure of *getfilaments* removes non-filamentary structures (noise and background fluctuations, sources), revealing clean filaments, such as the ones shown as (normalized) mask images $\mathcal{M}_{\lambda j}$ in Fig. 8. The masks define the maximum area of single-scale filaments at their base level, allowing reconstruction of their intensity distribution (Sect. 2.4.3).

The above method of detecting real filamentary structures in single-scale images produces very few (if any) spurious filaments and only for strongly-variable backgrounds. Even if a few spurious filaments are found, they are practically always quite faint and should not present any real problem since they can be easily removed, if necessary. Experience shows that the filament threshold $\tilde{\varpi}_{\lambda j} = \sigma_{\lambda j}$ is a good choice: decreasing it to even lower levels would result in more spurious filaments. Indeed, at progressively lower levels, small-scale noise or background fluctuations merge into longer, randomly-oriented elongated chains, similar to the white structures within the noise on the smallest spatial scales in Fig. 6. Some of them would have $N_{\Pi\lambda j} > N^{\min}_{\Pi\lambda j}$ and thus contaminate clean images of filaments with faint spurious structures.

After applying the cleaning procedure to $\mathcal{I}_{\lambda \mathrm{D}j}$ (Fig. 4), one can derive intensity distributions within the filament masks on each spatial scale (Fig. 9) and determine their full intensities by summing up all scales. However, one has to overcome several complications to properly reconstruct the intrinsic intensity distribution of filaments. On one hand, it is necessary to include negative areas of single-scale filaments (Fig. 4) resulting from the spatial decomposition of Eq. (2), as accumulating only positive intensities would incorrectly give substantially wider filaments. On the other hand, observed intensity distributions are very often altered by the bright sources spatially associated with filaments (Figs. 3–5). To include negative areas and remove practically entire contribution of sources, *getfilaments* reconstructs clean filament intensities $\tilde{\mathcal{I}}_{\lambda \mathrm{D}j\mathrm{C}}$ in a more elaborate way.

2.4.3. Reconstructing intrinsic intensities of filaments

The cleaning algorithm described in Sect. 2.4.2 detects all filamentary structures that exist in decomposed single-scale images $\mathcal{I}_{\lambda \mathrm{D}j}$, by separating the filamentary component of the images from all other structures of non-filamentary nature that are *outside* the filament masks. The problem is that the filaments in $\tilde{\mathcal{I}}_{\lambda \mathrm{D}j\mathrm{C}}$ are still contaminated by the noise and background fluctuations and by the sources that are *inside* their masks (Fig. 9). Substantial work is required to reconstruct full intrinsic intensity distribution of filaments, removing all non-filamentary peaks. In real-life observations, individual filaments and different segments of a filament usually have significantly varying intensities. This further complicates the problem, since *getsources* and *getfilaments* process entire images and neither individual filaments nor their parts.

Removing sources from filaments in each single scale, *getfilaments* splits images $\mathcal{I}_{\lambda \mathrm{D}j}$ between their maximum and $\tilde{\varpi}_{\lambda j}$ by a number of intensity levels l, spaced by a factor of 1.05. At each level, the filament reconstruction procedure works on a sequence of differential images

$$\delta\mathcal{I}^{+}_{\lambda \mathrm{D}jl} = \mathcal{I}_{\lambda \mathrm{D}jl+1} - \mathcal{I}_{\lambda \mathrm{D}jl} \tag{6}$$

from the bottom to the top, starting with $\tilde{\varpi}_{\lambda j}$ at $l = 1$. Clusters of connected pixels with $N_{\Pi\lambda j} < N^{\min}_{\Pi\lambda j}$ are removed from the images using the *TintFill* algorithm, producing clean images $\delta\mathcal{I}^{+}_{\lambda \mathrm{D}jl\mathrm{C}}$. This cleaning in the process of reconstruction of the intrinsic intensities of filaments is essentially the same as the procedure used above for obtaining clean masks of filaments, with the only difference being that the decomposition beam S_j in Eq. (4) is replaced by $\min\{S_j, 1.8\,O_\lambda\}$. The latter softens the degree of cleaning of the differential images on spatial scales larger than $3\,O_\lambda$, as the noise and background fluctuations and sources usually make smaller contribution to the filaments on large spatial scales. This softening of the cleaning allows an accurate reconstruction of the intrinsic filament profiles including their faintest outskirts on the largest scales.

The above approach worked very well in all benchmarks and real-life images where it has been tested. It was found to become less accurate only in a model image that simulated the limiting case of extremely large sources on top of comparably wide filaments, in which case the filament intensities under the source were overestimated and the source intensities were underestimated by up to ~30%. This is related to the fundamental difficulty of distinguishing between components on very large scales (approaching the image size), where signals from the components blend together so much that it is impossible to separate them without any additional assumptions. One may consider this case unrealistic, since all observed images tested were very far from displaying such combinations of sources and filaments.

Clean images $\delta\mathcal{I}_{\lambda \mathrm{D}jl\mathrm{C}}$ are summed up further with those from all lower levels. When all levels have been processed and accumulated, the resulting images (Fig. 10) contain reconstructed positive intensity distributions of all filamentary structures present in $\mathcal{I}_{\lambda \mathrm{D}j}$, with most of the peaks from noise,

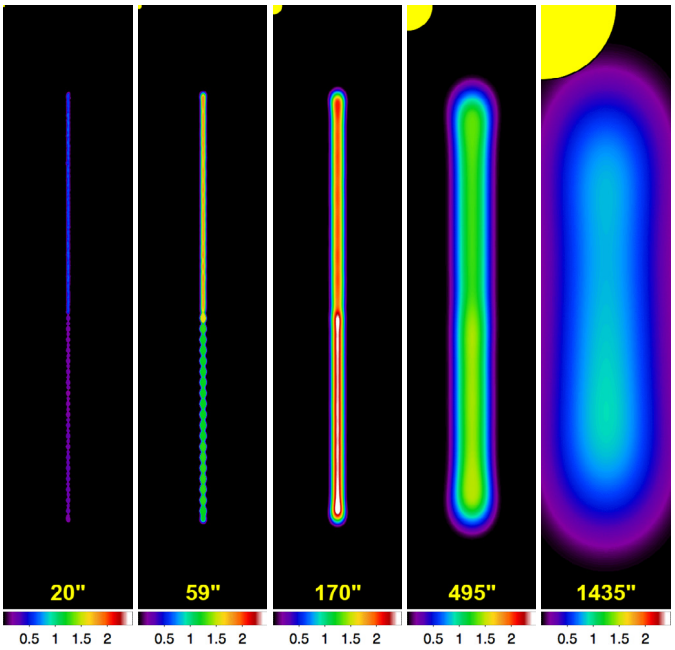

Fig. 10. Reconstructed positive areas in single-scale images of filaments (Sect. 2.4). The images of spatial scales from Fig. 4 are shown here in pixels with $I_{\lambda j} > \tilde{\varpi}_{\lambda j}$ after the removal of small structures (sources) from the filament (cf. Figs. 4, 9). This gives a relatively good approximation to its intrinsic intensity distribution of filaments; however, the images do not take the negative areas into account.

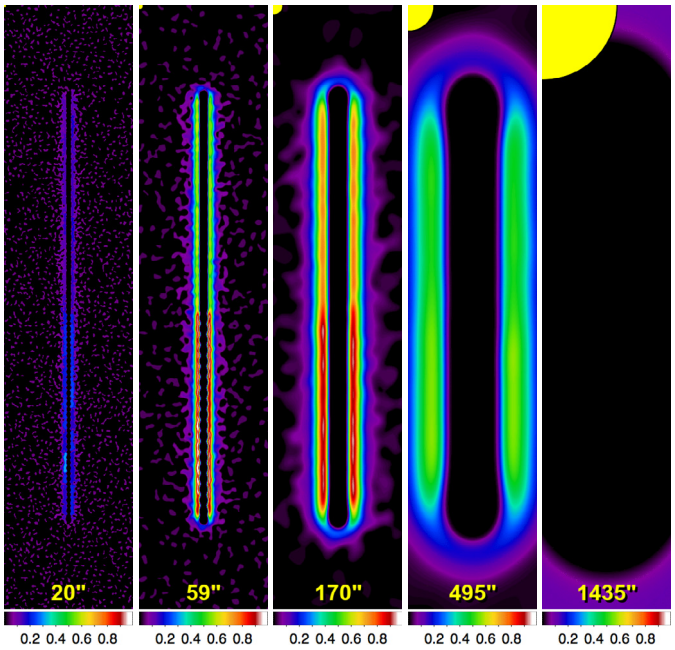

Fig. 11. Reconstructed negative areas in single-scale images of filaments (Sect. 2.4). The images of spatial scales from Fig. 4 are shown here in pixels with $I_{\lambda j} < -\tilde{\varpi}_{\lambda j}$ after the removal of small clusters with large negative values around bright non-filamentary peaks (sources) within the filament (cf. Figs. 4, 6). The images are shown here with positive values (multiplied by −1) for better visibility.

background, and sources removed:

$$\tilde{\mathcal{I}}^{+}_{\lambda \mathrm{D}j\mathrm{C}} = \sum_{l=0}^{N^{+}_{\mathrm{L}}} \delta\mathcal{I}^{+}_{\lambda \mathrm{D}jl\mathrm{C}}, \tag{7}$$

where N^{+}_{L} is the number of levels and the $l = 0$ base-level clean differential image is obtained from $\mathcal{I}_{\lambda \mathrm{D}j}$ by only taking pixels with $I_{\lambda j} \le \tilde{\varpi}_{\lambda j}$.

The negative areas around decomposed filaments also require a careful treatment. They cannot be taken directly from the decomposed images (Fig. 4), because the latter are non-locally affected by the negative areas produced by sources and other peaks of non-filamentary nature. To obtain the clean negative areas of the filaments, the algorithm multiplies the single-scale images by −1 and applies the same cleaning procedure described above for the positive areas. It splits the images $-\mathcal{I}_{\lambda \mathrm{D}j}$ between their maximum and $\tilde{\varpi}_{\lambda j}$ by a number of intensity levels

$$\delta\mathcal{I}^{-}_{\lambda \mathrm{D}jl} = -\left(\mathcal{I}_{\lambda \mathrm{D}jl+1} - \mathcal{I}_{\lambda \mathrm{D}jl}\right), \tag{8}$$

starting with $\tilde{\varpi}_{\lambda j}$ at $l = 1$, and produces clean images $\delta\mathcal{I}^{-}_{\lambda \mathrm{D}jl\mathrm{C}}$ by removing small clusters of connected pixels. All processed levels are accumulated, which gives reconstructed negative areas (Fig. 11) produced by all filamentary structures present in $\mathcal{I}_{\lambda \mathrm{D}j}$, with most of the negatives from noise, background, and sources removed:

$$\tilde{\mathcal{I}}^{-}_{\lambda \mathrm{D}j\mathrm{C}} = \sum_{l=0}^{N^{-}_{\mathrm{L}}} \delta\mathcal{I}^{-}_{\lambda \mathrm{D}jl\mathrm{C}}, \tag{9}$$

where N^{-}_{L} is the number of levels and the $l = 0$ base-level clean differential image is obtained from $\mathcal{I}^{-}_{\lambda \mathrm{D}j}$ by taking only pixels with $I_{\lambda j} \le \tilde{\varpi}_{\lambda j}$.

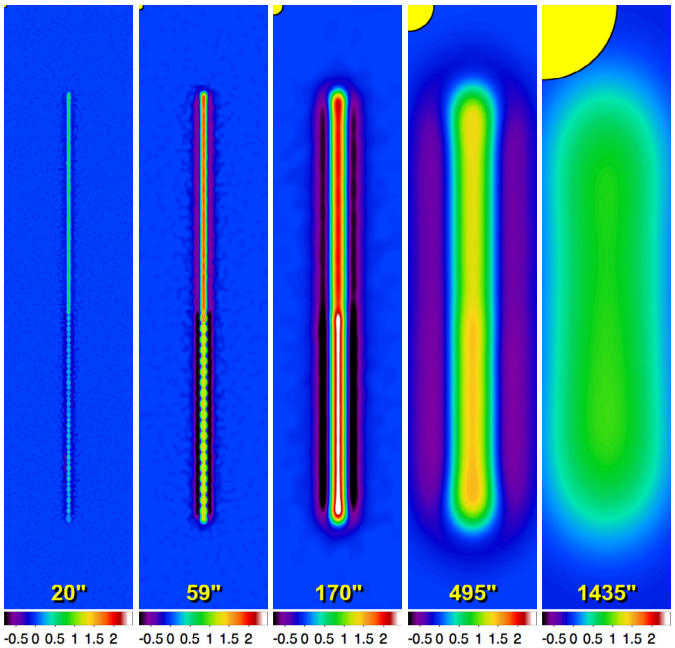

Fig. 12. Reconstructed clean single-scale filaments (Sect. 2.4). The images of spatial scales from Fig. 4 are shown here as $\tilde{\mathcal{I}}_{\lambda \mathrm{D}j\mathrm{C}}$ after the removal of noise and background fluctuations. Reconstructed positive and negative areas from Figs. 10 and 11 have been added together to produce the images.

The reconstructed positive and negative components are slightly convolved using a small Gaussian beam with a size of $0.1 \max\{S_j, \Delta\}$. The convolution is beneficial for avoiding abrupt intensity jumps to zero below $\tilde{\varpi}_{\lambda j}$ and it does not alter the filament intensity distribution above $\tilde{\varpi}_{\lambda j}$ because of the small beam. After computing the clean positives and negatives, one can easily

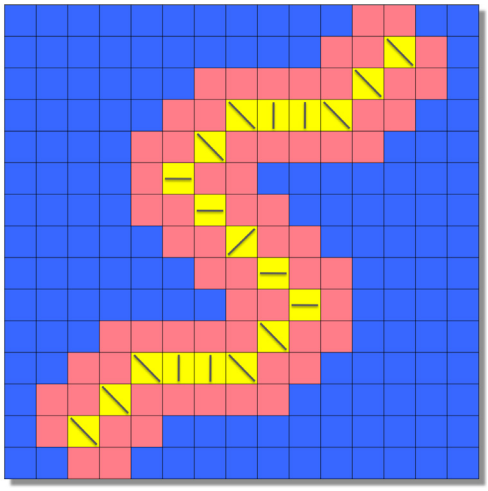

Fig. 13. Single-scale skeletons (Sect. 2.4). Red and yellow pixels belong to a clean reconstructed filament, whereas the blue pixels have been set to zero during the cleaning. The filament is measured at each pixel in the four main directions to determine the direction where the filament has the smallest half-maximum width. Along those directions (indicated by short straight lines), the brightest pixels (marked by yellow color) belong to the filament's crest and define the skeleton.

obtain their intensity distributions on each scale (Fig. 12):

$$\tilde{\mathcal{I}}_{\lambda \mathrm{D}j\mathrm{C}} \equiv \tilde{\mathcal{I}}^{+}_{\lambda \mathrm{D}j\mathrm{C}} - \tilde{\mathcal{I}}^{-}_{\lambda \mathrm{D}j\mathrm{C}}, \tag{10}$$

where all significant contributions of non-filamentary components have been removed. Reconstruction of the full intrinsic intensity distribution of all significant filaments detected in $\mathcal{I}_{\lambda \mathrm{D}}$ on all spatial scales reduces to

$$\tilde{\mathcal{I}}_{\lambda \mathrm{D}\mathrm{C}} = \sum_{j=1}^{N_\mathrm{S}} \tilde{\mathcal{I}}_{\lambda \mathrm{D}j\mathrm{C}} + \mathcal{G}_{N_\mathrm{S}} * \sum_{j=1}^{N_\mathrm{S}} \tilde{\mathcal{I}}_{\lambda \mathrm{D}j\mathrm{C}}, \tag{11}$$

where the second term is an estimate of the contribution of the clean filaments to the largest scales ($S > S_{\mathrm{max}}$) that are always left out of the single-scale decomposition (cf. Eqs. (2), (3)).

To improve the quality of measurements, image flattening, and final extraction (Sects. 2.6, 2.7), the full reconstructed filaments $\tilde{\mathcal{I}}_{\lambda \mathrm{D}\mathrm{C}}$ are subtracted from the original detection images $\mathcal{I}_{\lambda \mathrm{D}}$ and measurement images $\mathcal{I}_{\lambda \mathrm{O}}$, producing their filament-subtracted counterparts $\mathcal{I}_{\lambda \mathrm{D}\,\mathrm{FS}}$ and $\mathcal{I}_{\lambda \mathrm{O}\,\mathrm{FS}}$.

2.4.4. Deriving skeletons of filaments

For studying properties of filaments, it is useful to determine their skeletons, i.e., the lines of connected pixels tracing filaments' crests. While at the cleaning step, *getfilaments* creates single-scale skeletons $\tilde{\mathcal{S}}_{\lambda \mathrm{D}j\mathrm{C}}$ using a simple algorithm illustrated in Fig. 13. At each pixel of the reconstructed filament $\tilde{\mathcal{I}}_{\lambda \mathrm{D}j\mathrm{C}}$, half-maximum widths of the latter in four main directions (two image axes and two diagonals) are analyzed to find the direction of the narrowest intensity profile. The brightest pixel of the profile in that direction defines the skeleton pixel, the value of which is set to 1. Since the skeletons $\tilde{\mathcal{S}}_{\lambda \mathrm{D}j\mathrm{C}}$ are obtained independently on each scale, their location may somewhat fluctuate between scales. To reduce the influence of that uncertainty on accumulated skeletons, the skeleton width is increased to three pixels. This means that each pixel of the skeleton shown in Fig. 13 will eventually spread its value over all eight surrounding pixels.

Full skeletons $\tilde{\mathcal{S}}_{\lambda \mathrm{D}\mathrm{C}}$ accumulated over all scales are obtained by summation, similarly to Eq. (11). Such skeletons contain detailed information on their significance, as the value of any non-zero pixel is proportional to the number of spatial scales, where the pixel belongs to the skeleton. Using the same algorithm, *getfilaments* creates another version of skeletons, tracing the crests of the skeletons $\tilde{\mathcal{S}}_{\lambda \mathrm{D}\mathrm{C}}$, instead of those of the filaments $\tilde{\mathcal{I}}_{\lambda \mathrm{D}j\mathrm{C}}$. The skeletons $\tilde{\mathcal{S}}'_{\lambda \mathrm{D}\mathrm{C}}$ are one-pixel wide and their pixel values are equal to 1. They are further used to produce the segmentation images of filamentary structures, where all pixels belonging to a filament have the value of the filament's number (see, e.g., Figs. B.1f, C.1f in Appendices B, C).

2.5. Combining clean single scales over wavelengths

The cleaning algorithm outlined in Sect. 2.4 is applied to the single-scale detection images $\mathcal{I}_{\lambda \mathrm{D}j}$ independently for each wavelength λ. Combining information across several wavebands in the process of source extraction with *getsources* significantly improves the source detection and measurement qualities (Paper I). There are similar benefits of combining filaments obtained independently at each wavelength, because the robustness of the detection of filamentary structures (their significance) increases with the number of wavebands where the structures appear. Since only monochromatic images of the simulated filament are used in this paper, the combination of filaments over wavebands is not illustrated here.

The positive component $\tilde{\mathcal{I}}^{+}_{\lambda \mathrm{D}j\mathrm{C}}$ of filamentary structures, produced by *getfilaments* (Sect. 2.4), is subtracted from the clean decomposed images $\mathcal{I}_{\lambda \mathrm{D}j\mathrm{C}}$ created by *getsources* for each spatial scale, which improves the reliability of source detection in filamentary backgrounds. In practice, the subtraction is done just before *getsources* produces combined detection images $\mathcal{I}_{\mathrm{D}j\mathrm{C}}$ and $\mathcal{I}'_{\mathrm{D}j\mathrm{C}}$ (Sect. 2.4 in Paper I).

For an overview of filamentary structures in all wavebands in multi-wavelength extractions, *getfilaments* accumulates clean filaments $\tilde{\mathcal{I}}_{\lambda \mathrm{D}\mathrm{C}}$ over all detection wavelengths, creating a combined image $\tilde{\mathcal{I}}_{\mathrm{D}\mathrm{C}}$. Similarly, the images of skeletons $\tilde{\mathcal{S}}_{\lambda \mathrm{D}\mathrm{C}}$ are accumulated over all bands in a combined image $\tilde{\mathcal{S}}_{\mathrm{D}\mathrm{C}}$. Although no such combined images are directly involved in either source or filament extraction, they provide cumulative views of the filaments' appearance and their significance across wavelengths that are useful when studying filamentary structures.

2.6. Detecting, measuring, and visualizing sources

Source detection, measurements, and visualization in *getsources* (Fig. 1; Paper I) are practically unaffected by *getfilaments*, except that the *filament-subtracted* versions of the respective images are used to improve the source detection and measurement qualities. An additional benefit of the filament extraction is that linear scanning artifacts or radial spikes of the diffraction pattern that may be contaminating observed images are also detected as (spurious) filaments and removed.

Sources are detected in the filament-subtracted single-scale detection images $\mathcal{I}_{\mathrm{D}j\mathrm{C}\,\mathrm{FS}}$, which greatly reduces the chances of creating spurious sources in strongly filamentary backgrounds. Measurements of the sources' properties are also performed in the filament-subtracted images $\mathcal{I}_{\lambda \mathrm{O}\,\mathrm{FS}}$, which considerably improves the interpolation and subtraction of backgrounds, because the latter become largely isotropic. The visualization step produces a number of additional images, where the sources are overlaid on the images of filaments and skeletons, useful when

studying various objects and processes associated with filamentary structures (not only forming stars, but also galaxies or their clusters; cf. Appendices B, C).

2.7. Flattening background and noise fluctuations

The *Herschel* images of Galactic regions display highly variable backgrounds; standard deviations of the combined background and noise fluctuations (outside of sources) sometimes differ by orders of magnitude between various areas of a large image $\mathcal{I}_{\lambda D}$. Any global thresholding method would have difficulty handling such images, because the thresholds would not be equally good for all areas. This is why *getsources* employs a special approach (Fig. 1; Paper I): completing an *initial* source extraction, then flattening detection images based on the local intensity fluctuations outside sources, and then performing the *final* source extraction using flattened images.

With *getfilaments*, the flattening procedure for detection images remains essentially the same as described in Paper I, except that the scaling (flattening) image $\mathcal{I}_{\lambda F}$ is computed from the *filament-subtracted* detection images $\mathcal{I}_{\lambda D\,FS}$. The images $\mathcal{F}_\lambda$ now include the footprints of both sources and filaments, to avoid any possibility that imperfect extraction of filaments in the initial extraction would affect the accuracy of flattening. The flattened filament-subtracted images are used in *getsources* throughout the final extraction, replacing the original detection images $\mathcal{I}_{\lambda D}$ that were used in the initial extraction.

3. Extraction results

Results of both initial and final extractions of filaments and sources in the simulated image (Fig. 3e), used to illustrate *getfilaments* in this paper, are shown in Figs. 14, 15. The power-law filament is extracted quite well already after the *initial* run. Away from the ends of the filament, the reconstructed filament is slightly overestimated owing to imperfect cleaning and separation of the isotropic background (Fig. 14b). Closer to the ends of the filament, where the background intensity becomes lower, the filament is somewhat underestimated in the initial extraction. After the *final* run, the filament profile is very accurate, as is demonstrated by the profiles f_1, f_7 in Fig. 15. Although a small fraction of the background ended up in the filament, maximum deviations from the true model filament intensity distribution are still within 5% of its peak.

Only at the locations of sources in the original image, the deviations increase to about 20% levels, due to imperfect separation of the sources in the process of the filament reconstruction, as indicated by the profiles f_1–f_3 in Fig. 15. The accuracy level depends on the relative properties of sources, filaments, background, and noise. Maximum deviations are expected in the most difficult cases when the sources have size and brightness similar to those of the filament in which they are embedded, as in the present simulation. The more sources and filaments are dissimilar in their widths, the easier it is to separate them; the brighter a component is, the better it can be extracted and the less accurate the extraction of the fainter components.

Without any background (see below), the final extraction brings substantial improvements to the faint outskirts of the filament intensity profile. However, in the presence of the relatively bright background, it also leads to a slight increase in the background emission incorporated into the filament (Fig. 14c). There is a fundamental difficulty in separating contributions of structural components on the largest scales, since their contributions

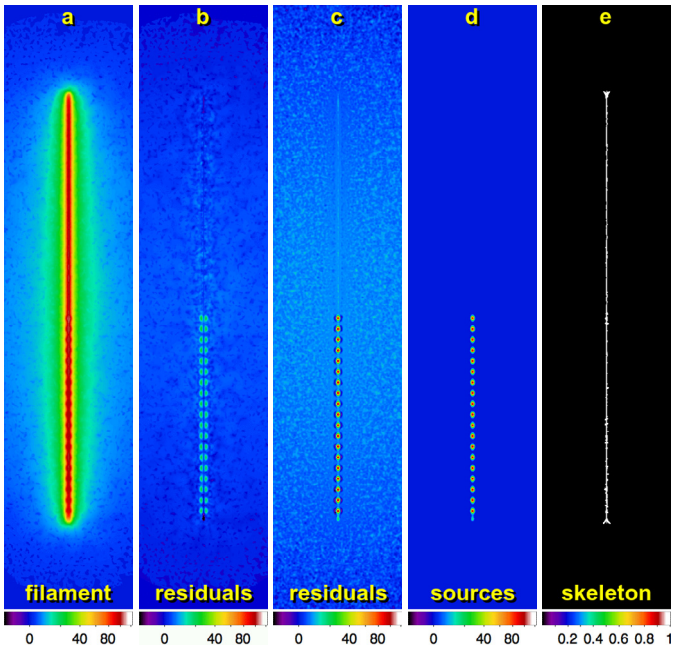

Fig. 14. Reconstruction of the simulated filament (Figs. 2, 3) obtained in the *final* extraction. **(a)** Intensity distribution of the clean reconstructed filament. **(b)** Residuals after the subtraction of the true model filament from the reconstructed filament. **(c)** Residuals after the subtraction of the reconstructed filament from the original simulated image. **(d)** Extracted sources, filament- and background-subtracted. **(e)** Skeleton, integrated over single scales, shown in those pixels that belong to it in more than 5 spatial scales.

blend together and become very similar. The power-law intensity profile $I_\lambda(r) \propto r^{-1}$ of the simulated filament makes the problem especially difficult. For steeper (e.g., Gaussian) filaments, the reconstruction and separation of all components are much more accurate.

Intensities of the extracted sources (Fig. 14d) are also somewhat affected by the process of separation: their profiles along the filament are slightly overestimated due to the emission from the filament, whereas in the orthogonal direction, they are somewhat underestimated (cf. profiles s_1, s_2 in Fig. 15). Despite the differences, the fluxes measured by *getsources* are sufficiently accurate, considering large *total* uncertainties involved in source extraction and measurements in highly structured and variable backgrounds[8]. The peak intensities of 19 sources were overestimated by ~1–7%, whereas their total fluxes were underestimated by ~10–20%. The variations in the accuracies of the fluxes are mainly caused by noise fluctuations, while their average levels are the consequence of the imperfection of the separation of the sources from the filament. Separating structural components from each other will always be a source of additional uncertainties, because the components are completely blended together, and their intrinsic intensity distributions are fundamentally unknown.

Extracting sources in filamentary images without first removing the filaments gives much less accurate results. To

[8] Total uncertainties include all possible sources: different sizes and complex shapes of observational beams, calibration, image reduction, map making, background subtraction, complex web of filamentary structures, noise and background fluctuations, dust opacities, optical depth effects, as well as the assumptions necessary to interpret the unknown three-dimensional reality on the basis of the observed two-dimensional images.

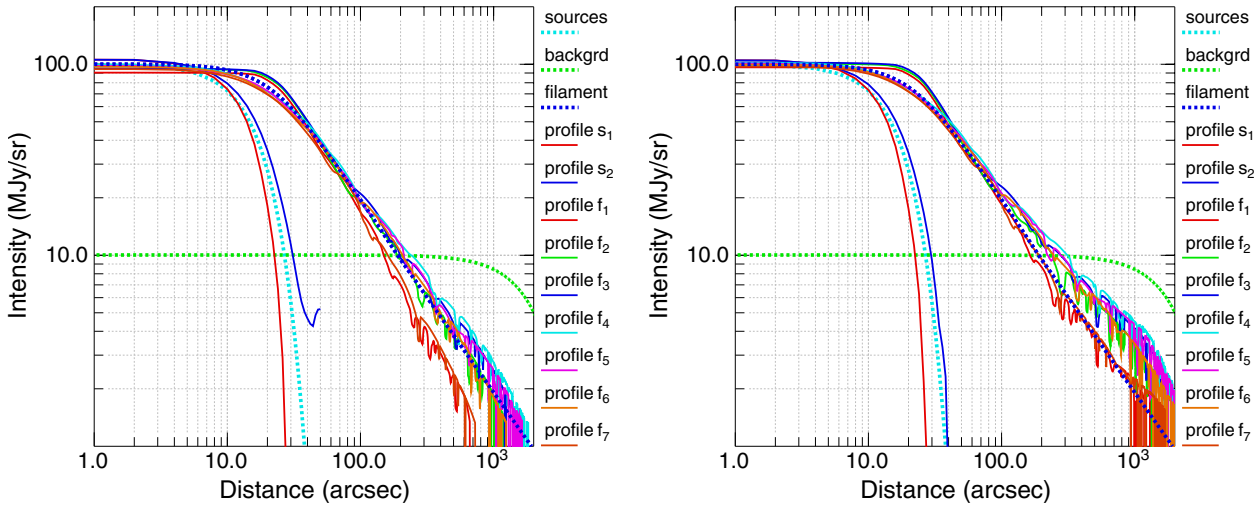

Fig. 15. Profiles of the components of the simulated image (Figs. 2, 3) obtained in the *initial* and *final* extractions (*left* and *right*, respectively). Orthogonal profiles of a model source (the second one from the bottom in Fig. 14d) are labeled s_1 and s_2; the narrower one (*red*) cuts through the source in a direction perpendicular to the filament, whereas the wider one (*blue*) is the source profile along the filament. Profiles f_1 to f_7 show cuts through the reconstructed filament (Fig. 14a) at 7 equidistant locations from its bottom to the top. The first three profiles (f_1–f_3) pass through the locations of the 2nd, 8th, and 14th sources; all other profiles (f_4–f_7) display the upper half of the filament, unaffected by sources.

demonstrate the difference, another *getsources* extraction was performed on the simulated image (Fig. 3e), with *getfilaments* turned off. On average, the sources were found to be substantially (by 60%) elongated along the filament. The peak intensities and total fluxes were overestimated by ∼30% and ∼100%, respectively. Local uncertainties of the fluxes were overestimated by almost an order of magnitude. The reason for the erroneous results is very simple: the sources were not separated from the filament and the fluxes and local intensity fluctuations include the signal from the filament. As is clear from Fig. 2, footprints of the model sources reach radial distances of 40″, where the filament intensity drops to 45 MJy/sr, by more than a factor of two. Assuming that the true source footprints are determined correctly, the background subtraction is to be done at that level of intensities, effectively incorporating the upper half of the filament into the sources. Along the filament, however, the baseline for background subtraction lies at 100 MJy/sr, the peak intensity of the filament. For such anisotropic "background" as the filament is, any approach to background subtraction that does not closely approximate the filament profile is bound to give inaccurate results.

The simulated filamentary image (Fig. 3e) was made relatively simple to demonstrate all features of the *getfilaments* method as clearly as possible; however, with four structural components it is not very simple (see much more complex cases in Appendices B, C). To illustrate the filament extraction with *getfilaments* in the most straightforward case, a simplified version of the filament image was created that combines just two of the components, the filament and noise (Fig. 3a, d). The intensity distribution and residuals of the extracted filament are displayed in Figs. 16 and 17. Only the results of the *final* extraction are presented because those of the *initial* run would appear almost indistinguishable. The small deficit visible in the reconstructed image and profile can be approximated by just a constant value of ∼1 MJy/sr. This minor discrepancy is caused by the difficulties in recovering the entire signal on the largest scales close to image size.

Compared to the simple simulation created for this paper, images obtained from complex three-dimensional simulations or *Herschel* observations are much more challenging for extracting and studying filamentary structures. Filaments in the interstellar

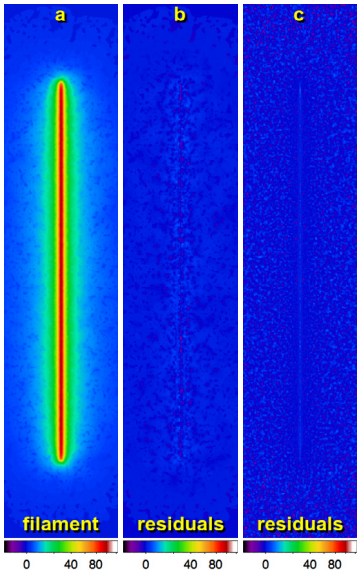

Fig. 16. Results of the *final* extraction for a simplified version of the simulated image (Figs. 2, 3) with only noise but *without* background and sources. **(a)** Intensity distribution of the clean reconstructed filament. **(b)** Residuals after the subtraction of the true model filament from the reconstructed filament. **(c)** Residuals after the subtraction of the reconstructed filament from the original simulated image.

medium appear to be very complex structures with greatly varying shapes, intensities, and profiles at different locations along their crests. Often they blend together with crowds of sources and with highly structured, bright, and variable backgrounds, as well as with other nearby filaments[9]. In multi-wavelength observations, the same filamentary area may appear quite differently owing to the contribution of structures with different temperatures and to optical depth effects. Unknown orientations of observed filaments in three-dimensional space greatly complicate their detailed studies and increase *total* uncertainties of results.

[9] Note that *getfilaments* does not attempt to *deblend* overlapping filaments, whereas *getsources* does deblending for overlapping sources: in contrast to filaments, sources can be reasonably approximated by simple deblending shapes.

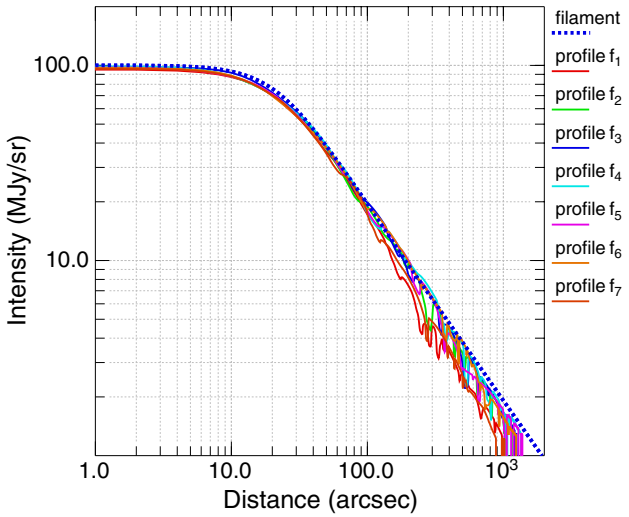

Fig. 17. Profiles of the reconstructed filament for a simplified version of the simulated image (Figs. 2, 3) with only noise but *without* background and sources, obtained in the *final* extraction. Profiles f_1 to f_7 show cuts through the extracted filament (Fig. 16a) at 7 equidistant locations from its bottom to the top.

4. Conclusions

Herschel observations have demonstrated that the interstellar medium is highly structured on all spatial scales and that its significant fraction emerges in omnipresent filamentary structures. Filamentary backgrounds present serious complications for source extraction methods since the filaments tend to amplify insignificant background or noise fluctuations that fall on top of the structures and thus create spurious sources. This paper describes the filament extraction method *getfilaments*, which shares the general multi-scale and multi-wavelength philosophy and approach with the source extraction method *getsources* (Paper I). Although both methods were designed primarily for use in large far-infrared and submillimeter surveys of star-forming regions with *Herschel*, they are applicable to other types of images.

Instead of tracing filaments directly in the observed images, *getfilaments* analyzes highly filtered decompositions of original images over a wide range of spatial scales (Sect. 2.3). The algorithm identifies filaments on each spatial scale as significantly elongated structures and reconstructs their full intrinsic intensity distributions, which are practically unaffected by sources and largely isotropic backgrounds (Sect. 2.4). Additionally, it determines single-scale and accumulated skeletons of the filaments, tracing the crests of their intensity distributions. Furthermore, it produces segmentation images of the filamentary structures, where each filament is identified by its sequential number. For an overview of all filaments, *getfilaments* creates combined images of clean filaments and their skeletons over all wavebands (Sect. 2.5). Based on the results of the initial extraction, detection images are flattened to produce much more uniform fluctuations of noise and non-filamentary background in preparation for the second, final extraction (Sect. 2.7).

Because it is incorporated into *getsources*, the *getfilaments* method brings substantial improvements to source extraction in filamentary backgrounds. Extraction of sources is also essential for an accurate reconstruction of the intrinsic intensity distribution of filaments. The intimate physical relationship between forming stars and filaments seen in *Herschel* observations demands that accurate filament extraction methods remove the contribution of sources, and conversely, accurate source extraction methods must be able to remove underlying filamentary structures. The images of clean filaments are now subtracted by *getsources* from the original images during source extraction, significantly improving the robustness of the method and reducing the chances of spurious sources contaminating extraction catalogs. An important benefit of the improved source extraction method is that, in addition to the catalogs and images of sources, it provides researchers with clean images of the filamentary structures that are the birthplace of stars.

Both *getsources* and *getfilaments* methods have been thoroughly tested using many simulated benchmark images and real-life observations. The source and filament extraction code is automated, very flexible, and easy-to-use. The latest version of the code with an installation guide and a quick start guide will soon be freely available upon request and downloadable from a web page[10].

Acknowledgements. The author employed *SAOImage DS9* (by William Joye) and *WCSTools* (by Jessica Mink) developed at the Smithsonian Astrophysical Observatory (USA), the *CFITSIO* library (by William D. Pence) developed at HEASARC NASA (USA), the *STILTS* library (by Mark Taylor) developed at Bristol University (UK), the *PSPLOT* library (by Kevin E. Kohler) developed at Nova Southeastern University Oceanographic Center (USA), *SWarp* (by Emmanuel Bertin) developed at Institut d'Astrophysique de Paris (France), and *ImageJ* (by Wayne Rasband) developed at the National Institutes of Health (USA). Web sites of the computational projects *Starformat* and *Horizon* provided images used for illustrations in this paper. Multi-wavelength observations of star-forming regions obtained in the frame of the *Herschel* Gould Belt and HOBYS Key Projects were useful in developing and testing both *getsources* and *getfilaments*; the work of all those who contributed to the production of the images (directly or indirectly) is highly appreciated. Useful tests were done on small ground-based (sub-) millimeter images of star-forming regions observed by Frédérique Motte and Philippe André. Two of a series of simulated filamentary images used in the development of *getfilaments* were created by Doris Arzoumanian, and some comparisons with observed filaments were made by Pedro Palmeirim. Comments on a draft made by Alana Rivera-Ingraham, Sarah Sadavoy, Nick Cox, Martin Hennemann, Nicola Schneider, Pierre Didelon, Nicolas Peretto, and Arabindo Roy are acknowledged.

References

André, P., Men'shchikov, A., Bontemps, S., et al. 2010, A&A, 518, L102
Arzoumanian, D., André, P., Didelon, P., et al. 2011, A&A, 529, L6
Bertin, E., Mellier, Y., Radovich, M., et al. 2002, in Astronomical Data Analysis Software and Systems XI, eds. D. A. Bohlender, D. Durand, & T. H. Handley, ASP Conf. Ser., 281, 228
Devriendt, J., Rimes, C., Pichon, C., et al. 2010, MNRAS, 403, L84
Hennebelle, P., Banerjee, R., Vázquez-Semadeni, E., Klessen, R. S., & Audit, E. 2008, A&A, 486, L43
Men'shchikov, A., André, P., Didelon, P., et al. 2010, A&A, 518, L103
Men'shchikov, A., André, P., Didelon, P., et al. 2012, A&A, 542, A81 (Paper I)
Moffat, A. F. J. 1969, A&A, 3, 455
Motte, F., Zavagno, A., Bontemps, S., et al. 2010, A&A, 518, L77
Ocvirk, P., Pichon, C., & Teyssier, R. 2008, MNRAS, 390, 1326
Palmeirim, P., André, P., Kirk, J., et al. 2013, A&A, 550, A38
Smith, A. R. 1979, in SIGGRAPH'79: Proc. of the 6th annual conference on Computer graphics and interactive techniques, 276 (New York: ACM)
Starck, J.-L., Elad, M., & Donoho, D. L. 2004, J. Advances in Imaging and Electron Physics, 132, 287

Pages 13 to 15 are available in the electronic edition of the journal at http://www.aanda.org

[10] http://gouldbelt-herschel.cea.fr/getsources

Appendix A: List of symbols

For the convenience of readers, this section lists and defines all symbols introduced in Sect. 2 of this paper (images are denoted by capital calligraphic characters):

$\mathcal{F}_\lambda$	images of source footprints in measurement iterations
$\mathcal{G}_j$	smoothing Gaussians in successive unsharp masking
$\mathcal{G}_\lambda$	smoothing Gaussians used to create detection images
$\mathcal{I}_{\mathrm{D}j\mathrm{C}}$	clean detection images combined over wavelengths
$\mathcal{I}_{\mathrm{D}j\mathrm{C}\,\mathrm{FS}}$	filament-subtracted combined detection images
$\mathcal{I}'_{\mathrm{D}j\mathrm{C}}$	clean detection images combined over wavelengths
$\mathcal{I}_\lambda$	original observed images produced by a map-maker
$\mathcal{I}_{\lambda\mathrm{DF}}$	flattened detection images for the final extraction
$\mathcal{I}_{\lambda\mathrm{D}}$	detection images: either $\mathcal{I}_{\lambda\mathrm{O}}$ or transformed $\mathcal{I}_{\lambda\mathrm{O}}$
$\mathcal{I}_{\lambda\mathrm{D}\,\mathrm{FS}}$	filament-subtracted detection images
$\mathcal{I}_{\lambda\mathrm{D}j}$	single-scale decompositions of the images $\mathcal{I}_{\lambda\mathrm{D}}$
$\mathcal{I}_{\lambda\mathrm{D}j\mathrm{C}}$	single-scale images cleaned of noise and background
$\tilde{\mathcal{I}}_{\lambda\mathrm{D}j\mathrm{C}}$	filaments cleaned of sources, noise, and background
$\tilde{\mathcal{I}}^{+}_{\lambda\mathrm{D}j\mathrm{C}}$	positive component of reconstructed filaments $\tilde{\mathcal{I}}_{\lambda\mathrm{D}j\mathrm{C}}$
$\tilde{\mathcal{I}}^{-}_{\lambda\mathrm{D}j\mathrm{C}}$	negative component of reconstructed filaments $\tilde{\mathcal{I}}_{\lambda\mathrm{D}j\mathrm{C}}$
$\mathcal{I}_{\lambda\mathrm{D}\,\mathrm{C}}$	full images of sources reconstructed from $\mathcal{I}_{\lambda\mathrm{D}j\mathrm{C}}$
$\tilde{\mathcal{I}}_{\mathrm{D}\,\mathrm{C}}$	image of filaments combined over wavelengths
$\tilde{\mathcal{I}}_{\lambda\mathrm{D}\,\mathrm{C}}$	full images of filaments reconstructed from $\mathcal{I}_{\lambda\mathrm{D}j}$
$\mathcal{I}_{\lambda\mathrm{F}}$	scaling image smoothed by convolution
$\mathcal{I}_{\lambda\mathrm{O}}$	measurement images: $\mathcal{I}_\lambda$ resampled to pixel Δ
$\mathcal{I}_{\lambda\mathrm{O}\,\mathrm{FS}}$	filament-subtracted measurement images
$\mathcal{M}_\lambda$	observational mask images defining areas of interest
$\mathcal{M}_{\lambda j}$	mask of a single-scale filament
$\tilde{\mathcal{S}}_{\mathrm{D}\,\mathrm{C}}$	image of skeletons combined over wavelengths
$\tilde{\mathcal{S}}_{\lambda\mathrm{D}j\mathrm{C}}$	skeletons of clean single-scale filaments
$\tilde{\mathcal{S}}_{\lambda\mathrm{D}\,\mathrm{C}}$	full accumulated skeletons of clean filaments
$\tilde{\mathcal{S}}'_{\mathrm{D}\,\mathrm{C}}$	wavelength-combined skeletons $\tilde{\mathcal{S}}'_{\lambda\mathrm{D}\,\mathrm{C}}$
$\tilde{\mathcal{S}}'_{\lambda\mathrm{D}\,\mathrm{C}}$	skeletons tracing crests of the full skeletons $\tilde{\mathcal{S}}_{\lambda\mathrm{D}\,\mathrm{C}}$
a	major size of a filament mask
A	major FWHM size of a source
$A^{\max}_\lambda$	maximum FWHM sizes of sources to be extracted
b	minor size of a filament mask
D_0	filament width: FWHM of the inner Gaussian core
$\tilde{E}$	elongation of the clusters of connected pixels
$\tilde{f}$	empirical shape factor of filamentary structures
f_{S}	scale factor defining relative spacing between scales
$f(\zeta)$	width normalization factor of a simulated filament
$I^*_{\lambda j}$	minimum peak intensity of detected filaments
$I_{\lambda j}$	pixel intensity in a single-scale detection image
$I_\lambda(r)$	intensity profile of a simulated filament
I_{P}	peak intensity of a simulated filament
j	running number of a decomposed spatial scale
l	running number of an intensity sub-level
L	length of a filament
$n_{\lambda j}$	variable number of standard deviations $\sigma_{\lambda j}$ in $\varpi_{\lambda j}$
N_{B}	number of cleaning beam areas
$N^{\pm}_{\mathrm{L}}$	number of intensity levels in filament reconstruction
N_{S}	number of spatial scales in the image decomposition
$N^{\min}_{\Pi\lambda j}$	minimum value of $N_{\Pi\lambda j}$ for cleaning filaments
$N_{\Pi\lambda}$	number of pixels in a cluster of connected pixels
O_λ	observational angular resolution: FWHM beam size
r	radial distance from the peak of a filament
R_0	filament radius: HWHM of the inner Gaussian core
S_j	spatial scale: FWHM of a smoothing Gaussian beam
$S_{\max}$	largest spatial scale in a single-scale decomposition
$\tilde{S}$	sparsity of the clusters of connected pixels
W	width of a filament
$\delta\mathcal{I}^{\pm}_{\lambda\mathrm{D}jl}$	differential images in filament reconstruction
Δ	pixel size (same for all images in an extraction)
λ	wavelength (central wavelength of a waveband)
$\varpi_{\lambda j}$	iterated cleaning thresholds (cut-off levels)
$\tilde{\varpi}_{\lambda j}$	filament detection thresholds (=$\sigma_{\lambda j}$)
$\sigma_{\lambda j}$	standard deviation in a single-scale image
σ_{noise}	standard deviation of simulated random noise
ζ	power-law exponent of a simulated filament

Appendix B: Filaments in MHD simulations

This section illustrates application of *getfilaments* to images obtained from three-dimensional magnetohydrodynamic (MHD) simulations of the formation of molecular clouds in colliding flows of warm diffuse gas (Hennebelle et al. 2008). Gravity, atomic cooling, photoelectric heating on dust grains, and initially uniform magnetic field were included in the simulations. Two opposite flows of diffuse neutral gas with the initial density of 1 cm^{-3} and velocity of 13.35 $\mathrm{km\,s}^{-1}$ were set up to collide in the *YZ* plane of the computational box. On a time scale of a few million years, a dense gas phase (10^2–10^4 cm^{-3}) developed under the influence of cooling, ram pressure, and gravity. All details of the simulation (labeled as *Slower Flow*) and corresponding images can be found on their web site[11].

A snapshot of the column density in the *YZ* plane corresponding to a time of 9.737 Myr from the start of the simulation was cut to a size of 1000×1000 pixels. The image was arbitrarily assigned a 2″ pixel size; the image values were scaled to a maximum of 100 (in arbitrary units) and some noise at a level of 0.5 has been added. The resulting image was convolved to a 5″ resolution.

The filamentary structures clearly visible in the original column density image (Fig. B.1a) are cleanly and almost fully extracted (Fig. B.1b), leaving only low-level filamentary residuals in the filament-subtracted image (Fig. B.1c). The latter shows mostly compact density enhancements (sources, intersections of the filaments) but no significant filamentary structures. An image of filaments reconstructed only up to a spatial scale of 20″ (Fig. B.1d) reveals the web of thin filaments that are largely diluted in panel b by the contribution of much larger scales. Although large filaments may appear as regular and smooth entities, many of them become heavily substructured on smaller scales. The composite image of the filaments (Fig. B.1e) uses the red, green, and blue colors to make the large, medium, and small-scale structures more visible. The segmented image of skeletons (Fig. B.1f) traces and numbers the crests of the filaments. All these images, as well as many other images and multi-wavelength catalogs of sources automatically produced by *getsources* and *getfilaments*, can be very useful for detailed studies of the properties of the filaments in the interstellar medium and their relationship with star formation.

Appendix C: Filaments in cosmological simulations

This section illustrates application of *getfilaments* to images obtained from the *Horizon MareNostrum* simulation of the formation of galaxies at high redshifts (Ocvirk et al. 2008; Devriendt et al. 2010) performed on the MareNostrum supercomputer at the Barcelona Supercomputer Center. Galactic winds, chemical enrichment, ultraviolet background heating, radiative cooling, star formation, and supernovae feedback were included in this

[11] http://starformat.obspm.fr/starformat/projects

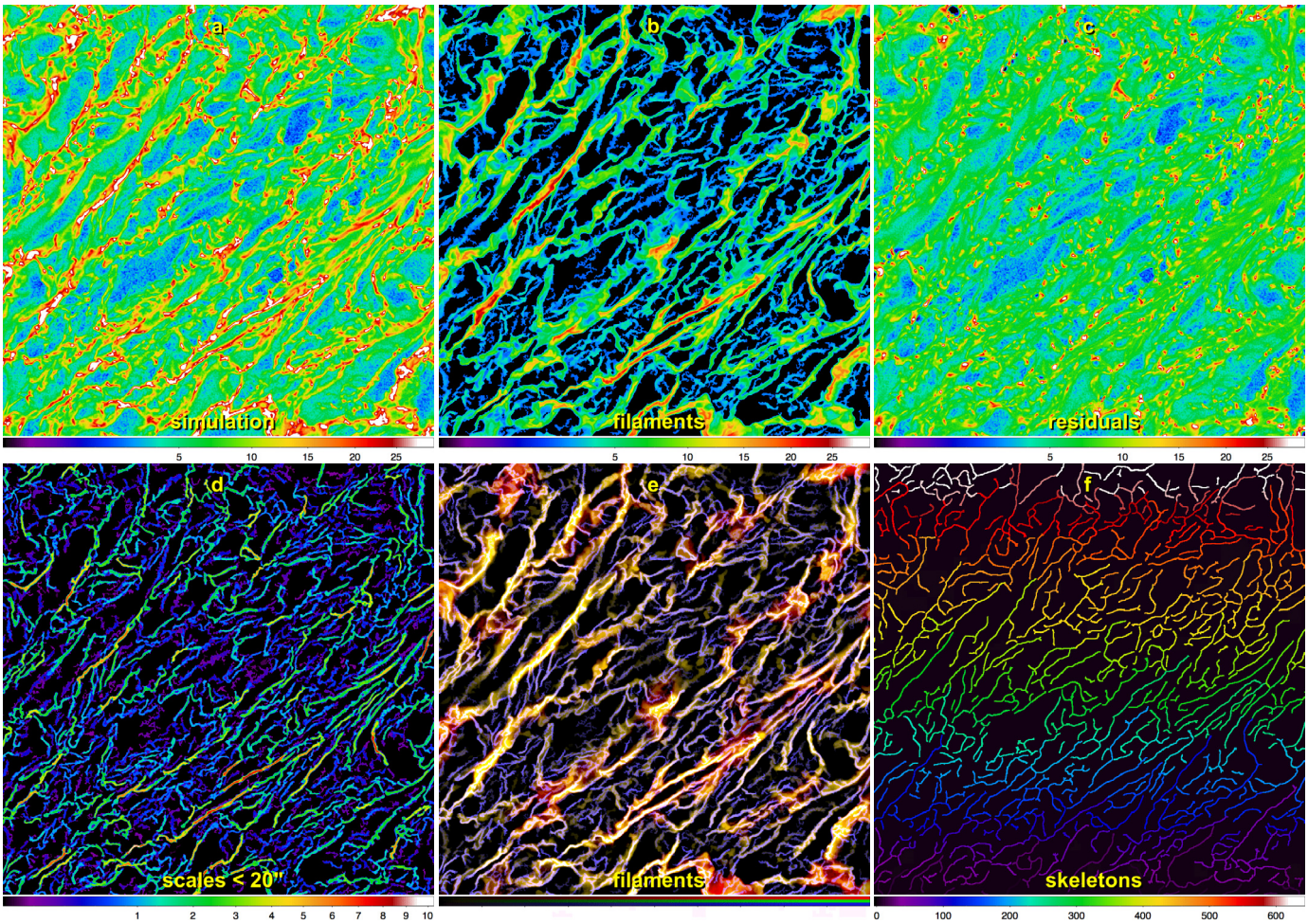

Fig. B.1. Filaments in MHD simulations of colliding flows (Hennebelle et al. 2008). The *upper panels* display the original image of column densities **(a)**, extracted filaments on all spatial scales **(b)**, and filament-subtracted image **(c)**. The *lower panels* show the filaments partially reconstructed up to 20″ scale **(d)**, 3-color (red, green, blue) composite image of the filaments partially reconstructed up to 2000″, 160″, and 10″ scales **(e)**, as well as the segmentation image of skeletons that appear on more than 5 spatial scales **(f)**. Pixel values in panel **(f)** represent the skeleton number.

large-scale and high-resolution simulation with up to five levels of adaptive mesh refinement. Impressive networks (cosmic web) of filamentary structures linking clusters of galaxies have been created and visualized in the simulation.

One of the images of a piece of the Universe corresponding to a redshift of 2.5 was downloaded from the project's web site[12], converted from JPG to FITS format using the *ImageMagick* utility, and reduced in size to 1000×1000 pixels. As in Appendix B, the image was arbitrarily assigned a 2″ pixel size, scaled to a maximum of 100 (in arbitrary units), and added with pixel noise at a level of 0.5. The resulting image was also convolved to a 5″ resolution.

The filament extraction results on cosmological scales are similar to those presented in Appendix B. The fascinating cosmic web visible in the original image (Fig. C.1a) is quite well extracted on all spatial scales (Fig. C.1b), with low filamentary residuals in the filament-subtracted image (Fig. C.1c) that shows mostly compact peaks (galaxies, clusters of galaxies). An image of filaments reconstructed up to a spatial scale of 20″ (Fig. C.1d) reveals thin filaments that are substantially diluted in panel b by the contribution of all larger scales; many large filaments are also substructured on smaller scales. The composite image of the filaments (Fig. C.1e) makes the large, medium, and small-scale structures more visible by combining the red, green, and blue colors on the same image. The segmented image of skeletons (Fig. C.1f) traces and numbers the crests of the filaments. Such images, as well as other images and multi-wavelength source catalogs produced by *getfilaments* and *getsources*, can readily be used for further studies of the cosmic web and the properties and formation processes of galaxies and their clusters.

[12] http://www.projet-horizon.fr

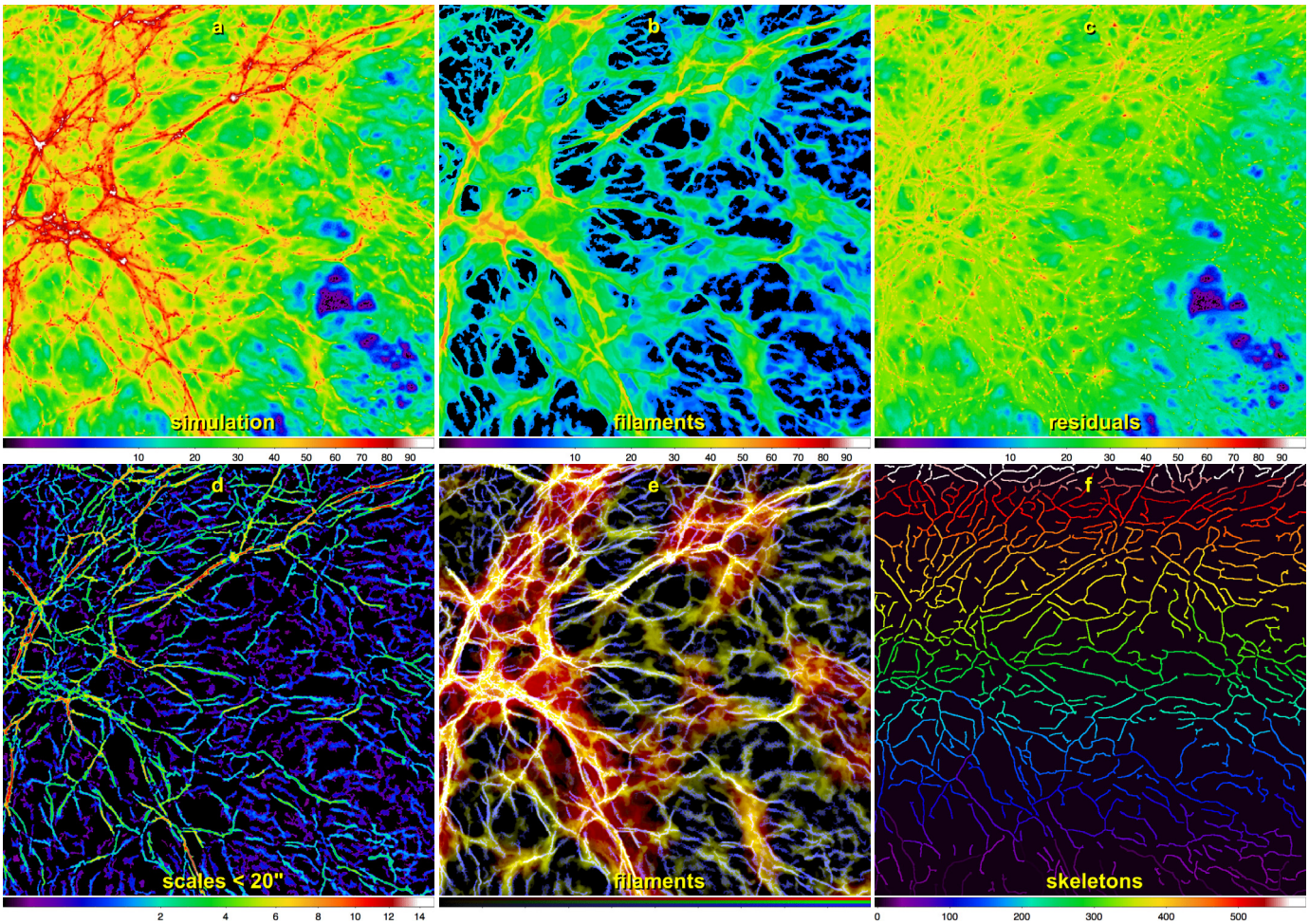

Fig. C.1. Filaments in the *MareNostrum* simulation of the formation of galaxies. The *upper panels* display the original image **(a)**, extracted filaments on all spatial scales **(b)**, and filament-subtracted image **(c)**. The *lower panels* show the filaments partially reconstructed up to 20″ scale **(d)**, 3-color (red, green, blue) composite image of the filaments partially reconstructed up to 2000″, 160″, and 10″ scales **(e)**, as well as the segmentation image of skeletons that appear on more than 5 spatial scales **(f)**. Pixel values in panel **(f)** represent the skeleton number.