STECKMAP: STEllar Content and Kinematics from high resolution galactic spectra via Maximum A Posteriori

P. Ocvirk, C. Pichon, A. Lançon, E. Thiébaut

Mon. Not. R. Astron. Soc. **365,** 74–84 (2006) doi:10.1111/j.1365-2966.2005.09323.x

STECKMAP: STEllar Content and Kinematics from high resolution galactic spectra via Maximum A Posteriori

P. Ocvirk,[1⋆] C. Pichon,[2] A. Lançon[1] and E. Thiébaut[3]

[1]*Observatoire de Strasbourg (UMR 7550), 11 Rue de l' Université, F-67000 Strasbourg, France*
[2]*Institut d' Astrophysique de Paris (UMR 7095), 98 bis Boulevard Arago, F-75014 Paris, France*
[3]*Observatoire de Lyon, 9 Avenue Charles André, F-69561 Saint Genis Laval, France*

Accepted 2005 June 14. Received 2005 May 10

ABSTRACT

We introduce STECKMAP (STEllar Content and Kinematics via Maximum A Posteriori likelihood), a method for recovering the kinematic properties of a galaxy simultaneously with its stellar content from integrated light spectra. It is an extension of STECMAP (presented recently by Ocvirk et al.) to the general case where the velocity distribution of the underlying stars is also unknown. The reconstructions of the stellar age distribution, the age–metallicity relation and the line-of-sight velocity distribution (LOSVD) are all non-parametric, i.e. no specific shape is assumed. The only a priori conditions that we use are positivity and the requirement that the solution is smooth enough. The smoothness parameter can be set by generalized cross-validation according to the level of noise in the data in order to avoid overinterpretation.

We use single stellar populations (SSPs) from PÉGASE-HR ($R = 10\,000$, $\lambda = 4\,000$–$6\,800$ Å, from Le Borgne et al.) to test the method through realistic simulations. Non-Gaussianities in LOSVDs are reliably recovered with signal-to-noise ratio (SNR) as low as 20 per 0.2 Å pixel. It turns out that the recovery of the stellar content is not degraded by the simultaneous recovery of the kinematic distribution, so that the resolution in age and error estimates given in Ocvirk et al. remain appropriate when used with STECKMAP.

We also explore the case of age-dependent kinematics (i.e. when each stellar component has its own LOSVD). We separate the bulge and disc components of an idealized simplified spiral galaxy in integrated light from high-quality pseudo-data (SNR = 100 per pixel, $R = 10\,000$), and constrain the kinematics (mean projected velocity, projected velocity dispersion) and age of both components.

Key words: methods: data analysis – methods: statistical – techniques: spectroscopic – galaxies: abundances – galaxies: kinematics and dynamics – galaxies: stellar content.

1 INTRODUCTION

For decades now, the spectral indices from the Lick group have been used to study the properties of stellar populations (Faber et al. 1985; Worthey 1994; Trager et al. 1998). Since the profile and depth of the lines involved in these spectral indices are affected by the line-of-sight velocity distribution (LOSVD) of the stars, it is necessary to correct the measured depths by a factor depending on the moments of the velocity distribution (Davies, Sadler & Peletier 1993; Kuntschner 2000, 2004). The latter moments must be determined using specialized code (Bender 1990; Kuijken & Merrifield 1993; van der Marel & Franx 1993; Saha & Williams 1994; Merritt 1997; Pinkney et al. 2003). These kinematic deconvolution routines have been used for some time and have undergone two major mutations. First, thanks to the increasing power of computers, it became affordable to swap back and forth from direct space to Fourier space, so that many disturbances such as border effects and saturation could be avoided. It became straightforward to mask problematic regions of the data, such as dead pixels, emission lines, etc. The second evolution of these codes allowed the use of multiple superimposed stellar templates to match best the observed spectrum (Rix & White 1992; Cappellari & Emsellem 2004). It has also been proposed to use single stellar populations (SSPs) as synthetic templates, and this approach has proved to be useful in addressing the template mismatch problem (Falcón-Barroso et al. 2003). Moreover, this technique can save precious telescope time, since it circumvents the need for observing template stars.

In Ocvirk et al. (2005, hereafter Paper I), we introduced STECMAP, a method for recovering non-parametrically the stellar

⋆ocvirk@pleiades5.u-strasbg.fr

content of a given galaxy from its integrated light spectrum. Using STECMAP requires, as a preliminary, convolving the data or models with the proper point spread function (PSF), which can be of both physical (i.e. the stellar LOSVD) and instrumental (the instrument's PSF) origin. Adjusting the LOSVD to fit the data not only constrains the kinematics of the observed galaxy but will also reduce the mismatch due to errors in the determination of the redshift or anomalous PSF, which is ultimately a necessary step when fitting galaxy spectra.

Here we propose to constrain the velocity distribution simultaneously with the stellar content, by merging the kinematic deconvolution and the stellar content reconstruction into one global maximum a posteriori likelihood inversion method. Hence, STECMAP becomes STECKMAP (STEllar Content and Kinematics via Maximum A Posteriori likelihood). In this respect, STECKMAP resembles the method proposed by, for example, Falcón-Barroso et al. (2003), except that it takes advantage of the treatment of the stellar content by STECMAP. Together with the stellar age distribution and the age–metallicity relation, the LOSVD is described nonparametrically and the only a priori conditions we use are smoothness and positivity.

We also tentatively address the case of age-dependent kinematics, i.e. we try to recover the individual LOSVDs and ages of several superimposed kinematic subcomponents. This approach is motivated by the fact that galaxies often display several kinematic components. Ellipticals and dwarf ellipticals, for instance, are known often to harbour kinematically decoupled cores (Balcells & Quinn 1990; Bender & Surma 1992; De Rijcke et al. 2004), and spiral galaxies are usually assumed to be constituted of a thin and a thick disc, a bulge and a halo (Freeman & Bland-Hawthorn 2002). The variety of the dynamical properties of the components has a counterpart in their stellar content, as a signature of the formation and evolution of the galaxy. For instance, the halo of the Milky Way is believed to consist mainly of old, metal-poor stars, while the bulge is more metal-rich, and the thin disc is mainly younger than the bulge (Freeman & Bland-Hawthorn 2002). It is thus natural to let any stellar subpopulation have its own LOSVD. This possibility has been recently addressed by De Bruyne et al. (2004a,b), in a slightly different framework. They use individual stars as templates for the different components, while we propose to use synthetic SSP models. Such a method would allow us to separate the several kinematic components of galaxies from integrated light spectra, and to constrain, for example, their age–velocity dispersion and age–metallicity relation. The highly detailed stellar content and kinematic information that can be obtained for the Milky Way or for nearby galaxies that can be resolved into stars, such as M31 (Ferguson et al. 2002; Ibata et al. 2004), could be extended to a larger sample of more distant galaxies. This technique could also be useful in detecting and characterizing major stellar streams in age and velocity from integral field spectroscopy of galaxies.

In this paper we use the PÉGASE-HR SSP models (Le Borgne et al. 2004) in order to illustrate and investigate the behaviour of the problems through simulations and inversions of mock data. Indeed, PÉGASE-HR, with its high spectral resolution ($R = 10\,000$), is an adequate choice for testing the recovery of detailed kinematic information in the form of non-parametric LOSVDs. The problems and methods we describe are, however, by no means specific to PÉGASE-HR (and its wavelength coverage), and STECKMAP could be used with any possible SSP model, depending on the type of data that is being analysed.

We will start with the modelling of the kinematics. Then, we will address the idealized linear problem of recovering the LOSVD when the stellar content is known, i.e. the template is assumed to be perfect. Section 4 deals with simultaneous age and LOSVD reconstruction of composite populations. Finally, Section 5 investigates the case of age-dependent kinematics in a simplified context where the metallicity and extinction are known a priori.

2 MODELS OF GALAXY SPECTRA

In this section we present the modelling of galaxy spectra, taking into account the composite nature of the stellar population, in age, metallicity and extinction, and finally its kinematics.

2.1 The composite reddened population at rest

We model the spectral energy distribution (SED) of the composite reddened population at rest using the ingredients defined in Paper I:

$$F_{\rm rest}(\lambda) = f_{\rm ext}(E,\lambda)\int_{t_{\rm min}}^{t_{\rm max}} \Lambda(t) B(\lambda, t, Z(t))\,{\rm d}t, \tag{1}$$

where $f_{\rm ext}(E,\,\lambda)$ is the extinction law, $\Lambda(t)$ is the luminosity-weighted stellar age distribution, $Z(t)$ is the age–metallicity relation, and $B(\lambda, t, Z)$ is the flux-averaged SSP basis of an isochrone population of wavelength λ, age t and metallicity Z. We recall briefly the main properties of the PÉGASE-HR SSP basis that we use in this paper. As mentioned earlier, spectral resolution is $R = 10\,000$ over the full optical domain $\lambda = [4000,\ 6800]$ Å, sampled in steps of 0.2 Å. The models are available for metallicities $Z \in [0.0001,\ 0.1]$ and considered reliable between $t_{\rm min} = 10$ Myr and $t_{\rm max} = 15$ Gyr. The initial mass function (IMF) used is described in Kroupa, Tout & Gilmore (1993), and the stellar masses range from 0.1 $\rm M_\odot$ to 120 $\rm M_\odot$. The extinction law $f_{\rm ext}$ was taken from Calzetti (2001).

2.2 Model kinematics

Stellar motions in galaxies define a LOSVD corresponding to projected local velocity distributions integrated along the line of sight and across one resolved spatial element.

2.2.1 Global kinematics

The motion of the stars can to first approximation be accounted for by assuming that the velocities of all stars of all ages along the line of sight are taken from the same velocity distribution (hence 'global'). The model SED, $\phi(\lambda)$, is the convolution of the assumed normalized LOSVD, $g(v)$, defined for $v \in [v_{\rm min}, v_{\rm max}]$ with the model spectrum at rest $F_{\rm rest}(\lambda)$. The convolved spectrum $\phi(\lambda)$ reads

$$\phi(\lambda) = \int_{v_{\rm min}}^{v_{\rm max}} F_{\rm rest}\left(\frac{\lambda}{1+v/c}\right) g(v)\,\frac{{\rm d}v}{1+v/c}, \tag{2}$$

where c is the velocity of light. The above expression reads as a standard convolution

$$\widetilde{\phi}(w) = c\int_{u_{\rm min}}^{u_{\rm max}} \widetilde{F}(w-u)\widetilde{g}(u)\,{\rm d}u, \tag{3}$$

with the following reparametrization:

$$w \equiv \ln(\lambda), \qquad u \equiv \ln\left(1+\frac{v}{c}\right), \tag{4}$$

$$\widetilde{F}(w) \equiv F_{\rm rest}({\rm e}^w) = F_{\rm rest}(\lambda), \tag{5}$$

$$\widetilde{g}(u) \equiv g(c({\rm e}^u-1)) = g(v), \qquad \widetilde{\phi}(w) \equiv \phi({\rm e}^w) = \phi(\lambda), \tag{6}$$

$$u_{\rm min} = \ln\left(1+\frac{v_{\rm min}}{c}\right), \qquad u_{\rm max} = \ln\left(1+\frac{v_{\rm max}}{c}\right). \tag{7}$$

2.2.2 Age-dependent kinematics

We now allow the LOSVD to depend on the age of the stars. For simplicity, we consider here only unreddened monometallic populations, i.e. $f_{\rm ext}(E, \lambda) = 1$ and $Z(t) = Z_0$. We introduce the age–velocity distribution, $\Xi(v, t)$, defined in $[v_{\rm min}, v_{\rm max}] \times [t_{\rm min}, t_{\rm max}]$, which gives the contribution of stars of velocity and age in $[v, v + {\rm d}v] \times [t, t + {\rm d}t]$ to the total observed light. Thus, for a given age t, $\Xi(v, t)$ is the LOSVD of the SSP of age t. The age–velocity distribution, $\Xi(v, t)$, is related to the stellar age distribution, $\Lambda(t)$, by

$$\int_{v_{\rm min}}^{v_{\rm max}} \Xi(v, t)\, {\rm d}v = \Lambda(t). \tag{8}$$

The model spectrum of such a population thus reads

$$\phi(\lambda) = \int_{t_{\rm min}}^{t_{\rm max}} \int_{v_{\rm min}}^{v_{\rm max}} B\left(\frac{\lambda}{1 + v/c},\ t,\ Z_0\right) \Xi(v, t)\, \frac{{\rm d}v\, {\rm d}t}{1 + v/c}. \tag{9}$$

The above expression can be rewritten more conveniently as

$$\widetilde{\phi}(w) = c \int_{u_{\rm min}}^{u_{\rm max}} \int_{t_{\rm min}}^{t_{\rm max}} \widetilde{B}(w-u, t) \widetilde{\Xi}(u, t)\, {\rm d}t\, {\rm d}u, \tag{10}$$

using the same reparametrization as in Section 2.2.1 and

$$\widetilde{B}(w, t) \equiv B({\rm e}^w, t, Z_0) = B(\lambda, t, Z_0), \tag{11}$$

$$\widetilde{\Xi}(u, t) \equiv \Xi(c({\rm e}^u - 1), t) = \Xi(v, t). \tag{12}$$

In the rest of this paper, we will use exclusively the standard (i.e. reparametrized) convolutions as in equations (3) and (10). For readability, we will drop the superscript $\widetilde{\ }$ and set the speed of light to unity.

3 KINEMATIC DECONVOLUTION

Section 2.2.1 shows that, with proper reparametrization, the convolution of a model spectrum at rest, $F(w)$, with the stellar LOSVD, $g(u)$, reads as a standard convolution, given by equation (3). Finding the LOSVD when the observed spectrum, $\phi(w)$, and the template spectrum, $F(w)$, are given is a classical deconvolution problem. Our goal here is not to discuss the respective qualities of the many different methods available in the literature to solve this problem. Most rely on fitting the data while imposing some a priori constraint on the LOSVD, i.e. they provide maximum a posteriori (MAP) estimates of the LOSVD. Let us describe briefly our method to obtain such a solution with the purpose of coupling it in a later step with STECMAP.

3.1 The convolution kernel

Here we discretize equation (3) to obtain a matrix form defining the convolution kernel. We use an evenly spaced set

$$\left\{u_j = u_{\rm min} + \left(j - \tfrac{1}{2}\right)\delta u; \quad j = 1, 2, \ldots, p\right\}$$

spanning $[u_{\rm min}, u_{\rm max}]$ with constant step $\delta u \equiv (u_{\rm max} - u_{\rm min})/p$. We expand the LOSVD as a sum of p gate functions:

$$g(u) = \frac{1}{\delta u} \sum_j g_j\, \theta\left(\frac{u - u_j}{\delta u}\right),$$

where

$$\theta(x) = \begin{cases} 1 & \text{if } -\frac{1}{2} < x \leqslant \frac{1}{2}, \\ 0 & \text{otherwise.} \end{cases}$$

Inserting this expansion into equation (3) leads to

$$\phi(w) = \frac{1}{\delta u} \sum_{j=1}^{j=p} g_j \int_{u_{\rm min}}^{u_{\rm max}} F(w - u)\, \theta\left(\frac{u - u_j}{\delta u}\right) {\rm d}u,$$

$$\simeq \sum_{j=1}^{j=p} g_j F(w - u_j). \tag{13}$$

Similarly, we now sample along the wavelengths by integrating over a small δw:

$$\phi_i \equiv \frac{1}{\delta w} \int \phi(w)\, \theta\left(\frac{w - w_i}{\delta w}\right) {\rm d}w,$$

$$\simeq \sum_{j=1}^{j=p} g_j F(w_i - u_j), \tag{14}$$

where $\{w_j\ ;\ j = 1, 2, \ldots, m\}$ is a set of *logarithmic* wavelengths spanning the spectral range with a constant step.

Using matrix notation and accounting for data noise, the observed SED reads

$$\boldsymbol{y} = \mathbf{K} \cdot \boldsymbol{g} + \boldsymbol{e}, \tag{15}$$

where $\mathbf{y} = (\phi_1, \phi_2, \ldots, \phi_m)^{\rm T}$ is the measured spectrum, and $\boldsymbol{e} = (e_1, e_2, \ldots, e_m)^{\rm T}$ accounts for modelling errors and noise. The vector of sought parameters $\boldsymbol{g} = (g_1, g_2, \ldots, g_p)^{\rm T}$ is the discretized LOSVD. The vector $\boldsymbol{s} = \mathbf{K} \cdot \boldsymbol{g}$ is the *model* of the observed spectrum, and the matrix $\mathbf{K}$,

$$K_{ij} = F(w_i - u_j), \quad \forall\, (i, j) \in \{1, \ldots, m\} \times \{1, \ldots, p\}, \tag{16}$$

is called the convolution kernel.

The convolution theorem (Press et al. 2002) states that the Fourier transform of the convolution of two functions is equal to the frequency-wise product of the individual Fourier transforms of the two functions. Applying this theorem yields another equivalent expression for the model spectrum **s**:

$$\boldsymbol{s} = \mathcal{F}^{-1} \cdot {\rm diag}(\mathcal{F} \cdot \boldsymbol{F}) \cdot \mathcal{F} \cdot \boldsymbol{g}, \tag{17}$$

where $\mathcal{F}$ is the discrete Fourier operator defined in Press et al. (2002) as

$$\mathcal{F}_{ij} = \exp\left[\frac{2{\rm i}\pi}{m}(i - 1)(j - 1)\right], \quad \forall\, (i, j) \in [1, \ldots m]^2, \tag{18}$$

$$\mathcal{F}^{-1} = \frac{1}{m}\mathcal{F}^*. \tag{19}$$

Note that, since m is the size of the template spectrum $\boldsymbol{F}$, the discretized LOSVD $\boldsymbol{g}$, which is initially of size p, needs to be symmetrically padded with zeros to the size m in order to transform the Toeplitz matrix into a circulant one. The diagonal matrix ${\rm diag}(\mathcal{F} \cdot \boldsymbol{F})$ carries the coefficients of the Fourier transform of the model spectrum at rest, $\boldsymbol{F}$. This notation involving the Fourier operator, $\mathcal{F}$, will be very useful for a number of algebraic derivations in the rest of the paper. In practice, from a computational point of view, it is more efficient to implement any forward or inverse Fourier transform through a fast Fourier transform (FFT). Similarly, the product ${\rm diag}(\mathcal{F} \cdot \boldsymbol{F}) \cdot \mathcal{F} \cdot \boldsymbol{g}$ is in practice implemented as a frequency-wise product of the individual FFTs.

3.2 Regularization and MAP

A number of earlier publications have shown that the maximum-likelihood solution to equation (15) is very sensitive to the noise in the data, $\boldsymbol{e}$. Hence, in the spirit of Paper I, we choose to regularize

the problem by requiring the LOSVD to be smooth. To do so, we use the quadratic penalization $P(\boldsymbol{g})$ as defined by equation (29) in Paper I:

$$P(\boldsymbol{g}) = \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{L}^{\mathrm{T}} \cdot \mathbf{L} \cdot \boldsymbol{g}. \quad (20)$$

In the rest of the paper, the penalization is Laplacian, i.e. $\mathbf{L} = \mathbf{D}_2$, where $\mathbf{D}_2$ is the discrete second-order difference operator, as defined in Pichon, Siebert & Bienaymé (2002). The objective function, Q_μ, to be minimized is given by

$$Q_\mu(\boldsymbol{g}) = \chi^2(\boldsymbol{y} \mid \boldsymbol{g}) + \mu P(\boldsymbol{g}), \quad (21)$$

where the χ^2 is defined by

$$\chi^2(\boldsymbol{y} \mid \boldsymbol{g}) = (\boldsymbol{y} - \boldsymbol{s}(\boldsymbol{g}))^{\mathrm{T}} \cdot \mathbf{W} \cdot (\boldsymbol{y} - \boldsymbol{s}(\boldsymbol{g})). \quad (22)$$

The vector $\mathbf{y}$ is the observed spectrum, and the weight matrix is the inverse of the covariance matrix of the noise: $\mathbf{W} = \mathrm{Cov}(\boldsymbol{e})^{-1}$. The parameter μ controls the smoothness of the LOSVD through its coefficients, $\boldsymbol{g}$. It can be set on the basis of simulations (as described in Paper I) or automatically by generalized cross validation (GCV) (Wahba 1990), according to the signal-to-noise ratio (SNR) of the data. In the latter case, the properties of the convolution kernel can be used to speed up the computation of the GCV function. Further regularization is provided by the requirement of positivity, implemented through quadratic reparametrization. Minimizing Q_μ yields the regularized solution $\boldsymbol{g}_\mu$. Efficient minimization procedures require the analytical expression of the gradients of Q_μ, given in Appendix Section A1.

3.3 Simulations

We applied this deconvolution technique to mock data, created from PÉGASE-HR SSPs of several ages and metallicities, with $R = 10\,000$ at 4000–6800 Å. In a first set of experiments, the model spectrum at rest was a solar-metallicity 10-Gyr SSP. It was convolved with various LOSVDs, both Gaussian and non-Gaussian, with velocity dispersions ranging from 30 to 500 km s^{-1}. It was then perturbed with Gaussian noise at levels ranging from SNR = 5 to 100 per pixel, and deconvolved using the model spectrum at rest as template (i.e. no template mismatch). In all cases, the LOSVDs are adequately recovered. Fig. 1 shows the reconstruction of a Gaussian LOSVD, for SNR = 10 per pixel. However, there are necessarily some biases in the reconstruction of the sharp features of the LOSVD. This is expected since we introduced regularization via smoothing. To illustrate the relationship between regularization and bias, we performed a new set of similar simulations for a non-Gaussian LOSVD (sum of two Gaussians) with SNR = 20 per pixel and varied the smoothing parameter μ. The results are shown in Fig. 2. Panels (a) and (b) correspond to $\mu = 10$, while panels (c) and (d) correspond to $\mu = 1000$. The model, median and interquartiles of 500 reconstructions are displayed. We also plotted the whole set of 500 recovered solutions, in order to show the locus of the solutions. One can see that the biases of the median reconstruction are reduced when lowering μ. The highest bump is correctly reproduced for $\mu = 10$, while it is not for $\mu = 1000$. On the other hand, the solutions are much more widely spread when $\mu = 10$. This means that most solutions taken from the set of low-μ simulations can be very far from the model, while all the large-μ solutions lie reasonably close to the model.

The regularization acts as a Wiener filter in the sense that it damps the high-frequency components of the solution. Regularization improves the significance of an individual reconstruction (it will nearly always lie reasonably close to the model), at the cost of introducing a bias.

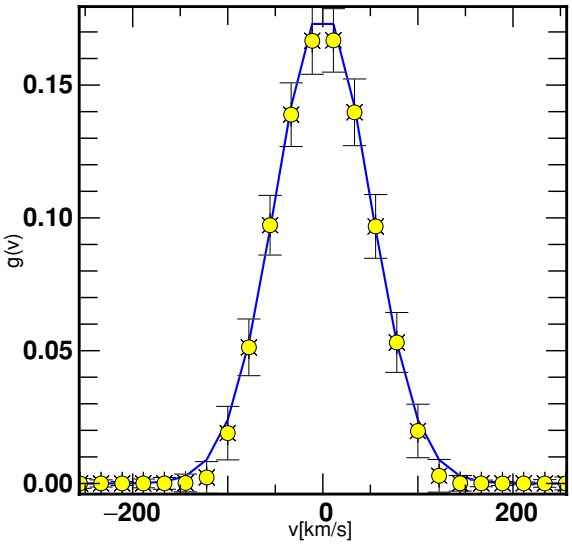

Figure 1. Non-parametric reconstruction of a Gaussian LOSVD for simulated data, $\sigma_v = 100$ km s^{-1}, SNR = 10 per pixel. The model spectrum at rest is a 10 Gyr old solar-metallicity SSP with $R = 10\,000$ at 4000–6800 Å. The template spectrum is identical, so that no template mismatch is allowed here. The curve is the input model. The circles and the bars show respectively the median and the interquartiles of the recovered solutions for 500 realizations of the noise.

3.4 Age and metallicity mismatch

We take advantage of the large range of ages and metallicities of SSPs covered by PÉGASE-HR to illustrate briefly the effects of template mismatch on LOSVD determinations. In this section we show the results of a large number of simulations aimed at characterizing the error made when a wrong template is chosen for the kinematic inversion of data. For this purpose, mock data were created by convolving an SSP of age a_0 and metallicity Z_0 with a centred Gaussian LOSVD of dispersion $\sigma_v = 50$ km s^{-1}. It was perturbed by Gaussian noise corresponding to SNR = 100 per pixel and then deconvolved, using as template an SSP of age a_1 and metallicity Z_1. The spectral resolution and wavelength range are the same as in Section 3.3. Fig. 3 shows the error on the measured velocity dispersion. The latter is measured as the rms of the reconstructed LOSVD. If the parameters of the template are different from those of the model, the velocity dispersion error increases very quickly. The age–metallicity degeneracy is visible as a valley of smaller error, following the upper left to bottom right diagonal of the figures. Of course, the χ^2 distance between the model and the mock data follows a similar 2D distribution, and will lead to the rejection of highly mismatched LOSVD estimates. However, in practice, it is usually not straightforward to quantify all the sources of error. It is thus somewhat arbitrary to set an upper limit of χ^2 for the admissible solutions, and the error on the kinematics is thus hard to quantify. This experiment illustrates in this context the long known issue that, when the correct model is not available, large errors on the determination of kinematics are expected. In order to reduce the error in the estimates of the kinematic properties of a stellar assembly, it is necessary to allow for a wide range of modulations of the template. This is naturally achieved by making the non-parametric stellar content account for the changes of the template, as discussed in the next section.

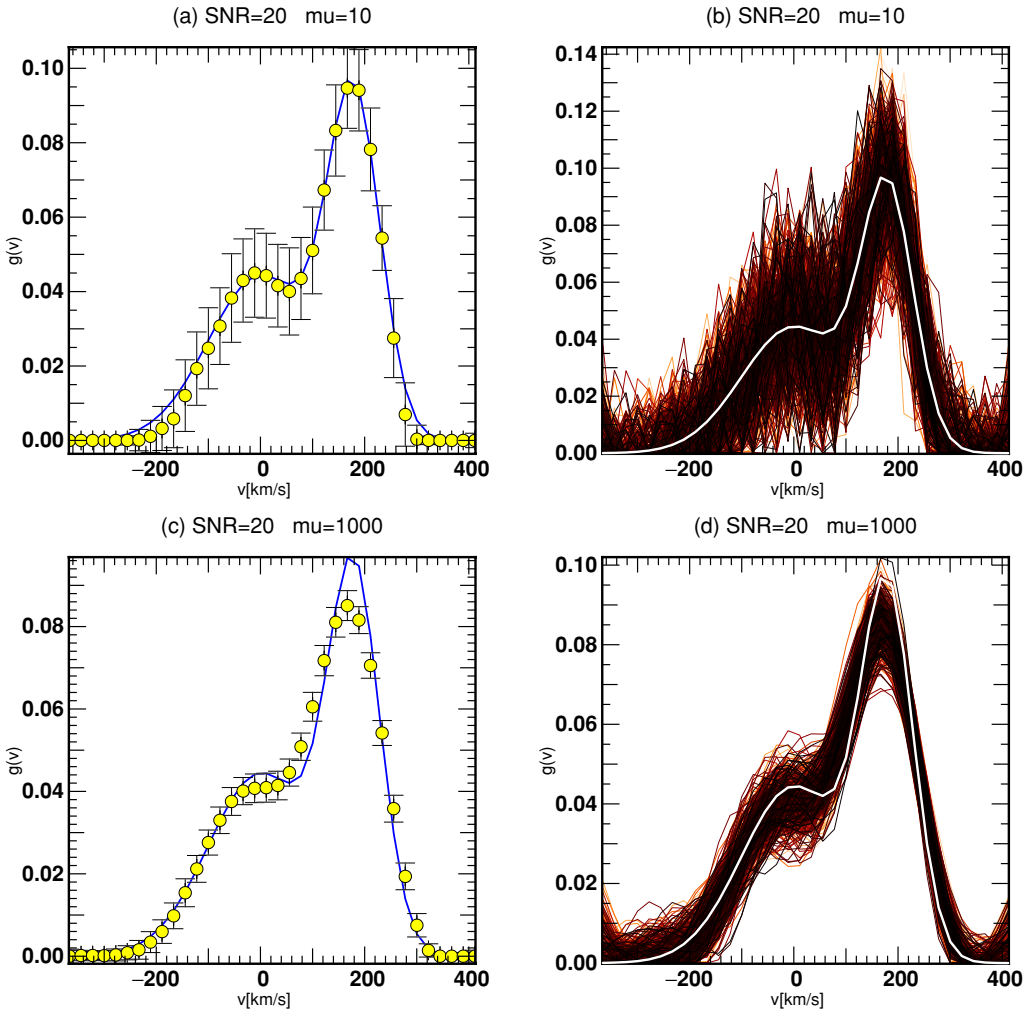

Figure 2. Impact of the smoothing parameter. Reconstruction of a non-Gaussian LOSVD for simulated data with SNR = 20 per pixel for $\mu = 10$ (top) and $\mu = 1000$ (bottom). Left: The curve is the input model LOSVD. The circles and the bars show respectively the median and interquartiles of the reconstructed LOSVDs for 500 realizations of the experiment. Right: The whole set of 500 solutions is displayed, with the model as a thick white line, in order to give the reader a sense of what individual solutions look like. The figures show the trade-off between bias and reliability of the reconstruction. For small μ, the median reconstruction is unbiased but the individual reconstructions are very noisy. For large μ, the median reconstruction is slightly biased but all the reconstructions are reasonable. Hence, the significance of an individual reconstruction is improved by regularization at the cost of introducing a bias.

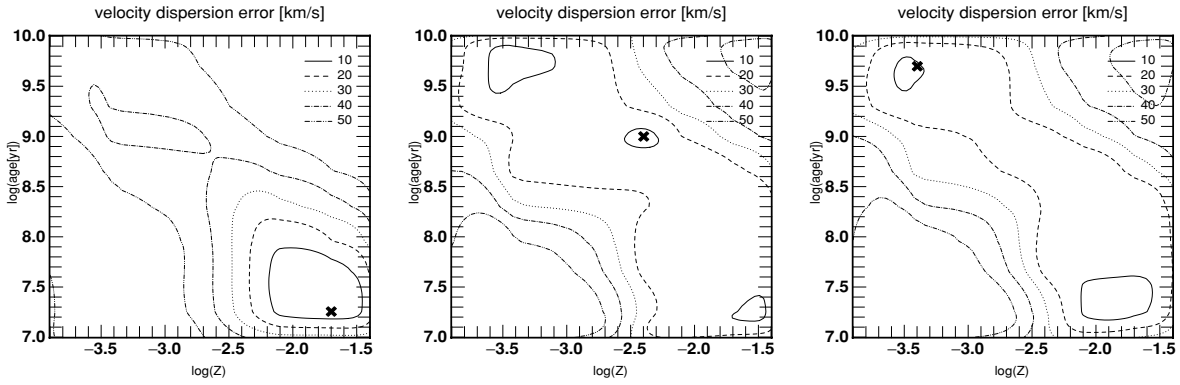

Figure 3. Velocity dispersion error as a function of the age and metallicity of the template SSP. Contours show regions of increasing velocity dispersion error. In each experiment, the age and metallicity of the original model template are shown as a thick cross, and the model LOSVD is a Gaussian with zero mean and 50 km s^{-1} dispersion. The velocity dispersion error is minimum when the template' s age and metallicity are similar to those of the model. The error increases quickly when the template parameters differ from the model parameters, also in the age–metallicity degeneracy direction (upper left to bottom right diagonal). It increases even faster in the direction orthogonal to the age–metallicity degeneracy.

4 RECOVERING STELLAR CONTENT AND GLOBAL KINEMATICS

The mixed inversion described in this section couples the recovery of both the stellar content and the kinematics, thereby turning STECMAP into STECKMAP. Proper application of this method provides an interpretation of the observed object in terms of stellar content and kinematics.

4.1 Inverse problem

For a given model spectrum at rest, $F_{\rm rest}(\lambda)$, and an LOSVD, $g(v)$, the emitted SED, $\phi(\lambda)$, is given by equation (2). We now wish to account also for the additional variables involved in $F_{\rm rest}$, given by equation (1), namely the stellar age distribution, $\Lambda(t)$, the age–metallicity relation, $Z(t)$, and the colour excess, $E(B-V)=E$. Inserting equation (1) into the convolution equation (3) yields the emitted SED

$$\phi(w) = \int\!\!\int f_{\rm ext}(E, w-u)\Lambda(t)B(w-u, t, Z(t))g(u)\,{\rm d}t\,{\rm d}u. \quad (23)$$

Solving equation (23) for Λ, Z, E and g when $\phi, f_{\rm ext}$ and B are given is the inverse problem we are tackling here.

4.2 Discretization and parameters

Expanding the two time-dependent unknowns $\Lambda(t)$ and $Z(t)$ as a sum of n gate functions and inserting into equation (1) yields the discrete model spectrum at rest:

$$\boldsymbol{F} = {\rm diag}(\boldsymbol{f}_{\rm ext}(E)) \cdot \mathbf{B} \cdot \boldsymbol{x}, \quad (24)$$

This discretization is explained in detail in section 5 of Paper I. Similarly, we develop the LOSVD $g(u)$ as a sum of p gate functions as in Section 3. Note that the reddened model at rest plays the role of the stellar template in a classical kinematic convolution. Inserting equation (24) into equation (17) thus allows us to express the model spectrum, $\boldsymbol{s}$, as

$$\boldsymbol{s} = \mathcal{F}^* \cdot {\rm diag}(\mathcal{F} \cdot {\rm diag}(\boldsymbol{f}_{\rm ext}(E)) \cdot \mathbf{B} \cdot \boldsymbol{x}) \cdot \mathcal{F} \cdot \boldsymbol{g}, \quad (25)$$

However, here, the template is this time modulated by the unknowns describing the stellar content.

4.3 Smoothness and metallicity constraints

The discrete problem of finding the stellar age distribution $\boldsymbol{x}$, the age–metallicity relation $\boldsymbol{Z}$, the extinction E and the LOSVD $\boldsymbol{g}$ for an observed spectrum $\boldsymbol{y}$ and given an extinction law $f_{\rm ext}$ and an SSP basis $\mathbf{B}$ is of course likely to be very ill-conditioned since it arises as the combination of several ill-conditioned problems. It therefore requires regularization. We also want the metallicity of the components to remain in the model range. We use the standard penalization P and the binding function C defined in Paper I to build the penalization P_μ for this problem:

$$P_\mu = \mu_x P(\boldsymbol{x}) + \mu_Z P(\boldsymbol{Z}) + \mu_C C(\boldsymbol{Z}) + \mu_v P(\boldsymbol{g}), \quad (26)$$

where $\boldsymbol{\mu} \equiv (\mu_x, \mu_Z, \mu_C, \mu_v)$. Again, we choose $\mathbf{L} = \mathbf{D}_2$ as defined in Pichon et al. (2002), so that the penalization P is actually Laplacian. The objective function, Q_μ, is now defined as

$$Q_\mu = \chi^2(\boldsymbol{s}(\boldsymbol{x}, \boldsymbol{Z}, E, \boldsymbol{g})) + P_\mu(\boldsymbol{x}, \boldsymbol{Z}, E, \boldsymbol{g}), \quad (27)$$

and its partial derivatives are given in Appendix Section A2. Note that there is in principle an additional formal degeneracy for this inverse problem. If the set $(\boldsymbol{x}, \boldsymbol{Z}, E, \boldsymbol{g})$ is a solution to (23), then $(\alpha\boldsymbol{x}, \boldsymbol{Z}, E, \boldsymbol{g}/\alpha)$ is also a solution for any scalar α, because the age distribution $\boldsymbol{x}$ and the LOSVD $\boldsymbol{g}$ are not explicitly normalized in this formulation. However, the adopted regularization lifts this degeneracy. The penalization function P is quadratic $[P(\alpha\boldsymbol{x}) = \alpha^2 P(\boldsymbol{x})]$. Thus, if $\boldsymbol{x}$ or $\boldsymbol{g}$ is too large in norm, the solution is unattractive. Practically, the algorithm reaches a solution where $\boldsymbol{x}$ and $\boldsymbol{g}$ are similar in norm. In any case, this degeneracy would easily be remedied by adding a normalizing term to the penalization P_μ of the form $\|\boldsymbol{x}\| - 1$, which would force the discretized stellar age distribution $\boldsymbol{x}$ to have unitary norm. Following the same principle, one could equivalently choose to normalize the LOSVD rather than the stellar age distribution.

4.4 Simulations

Let us now test the behaviour of STECKMAP by applying it to mock data. The latter were produced using an arbitrary stellar age distribution $\boldsymbol{x}$, an age–metallicity relation $\boldsymbol{Z}$, an LOSVD $\boldsymbol{g}$ and an extinction parameter E. Several simulations were performed with various input models: bumpy age distributions, increasing or decreasing age–metallicity relation and extinctions, Gaussian and non-Gaussian wide or narrow LOSVDs, in various pseudo-observational contexts. Fig. 4 shows the results of two of these experiments. In the top line, the model is a young metal-poor population superimposed on to an older metal-rich population. In the bottom panels, the model has a rather constant stellar age distribution, a non-monotonic age–metallicity relation and a strongly non-Gaussian LOSVD. In both cases the three unknowns are correctly recovered. In these examples, the data quality mimics that of the best Sloan Digital Sky Survey galaxies: the resolution is $R \approx 2000$ and SNR = 30 per ≈ 1 Å pixel. The wavelength domain of PÉGASE-HR is however narrower than that of the SDSS. These simulations simply aim to demonstrate the generally good behaviour of the method, and show that accounting for the kinematics does not fundamentally weaken the constraints on the stellar content. For a more thorough study of the informational content of the PÉGASE-HR wavelength range, the reader can refer to the systematic double burst simulations with variable spectral resolution and SNR per Å performed in Paper I.

5 RECOVERY OF AGE-DEPENDENT KINEMATICS

In this section we present an implementation of the recovery of age-dependent kinematics, i.e. the situation when each subpopulation has its *own* LOSVD. In this experiment, we restrict ourselves to the case where the stellar populations have a known metallicity and are seen without extinction. This choice is mainly motivated by the numerical cost of such a large inversion procedure. The modelling is given by equation (10). Finding the age–velocity distribution $\Xi(u, t)$ when the monometallic basis $\mathbf{B}$ and the observed spectrum ϕ are given is the inverse problem. It arises as the combination of a linear age inversion and a kinematic deconvolution.

5.1 A sum of convolutions

The age–velocity distribution, $\Xi(u, t)$, is expanded as a linear combination of normalized 2D gate functions $\theta_{ij}(u, t)$:

$$\theta_{ij}(u, t) \equiv \frac{1}{\delta u\,\delta t}\,\theta\left(\frac{u-u_i}{\delta u}\right)\theta\left(\frac{t-t_j}{\delta t}\right).$$

In other words, $\Xi(u, t)$ is represented by a 2D array $\mathbf{v}$ of size (p, n), i.e. p is the size of each LOSVD and n is the number of age bins. The linear step in u is δu and the step in t is δt.

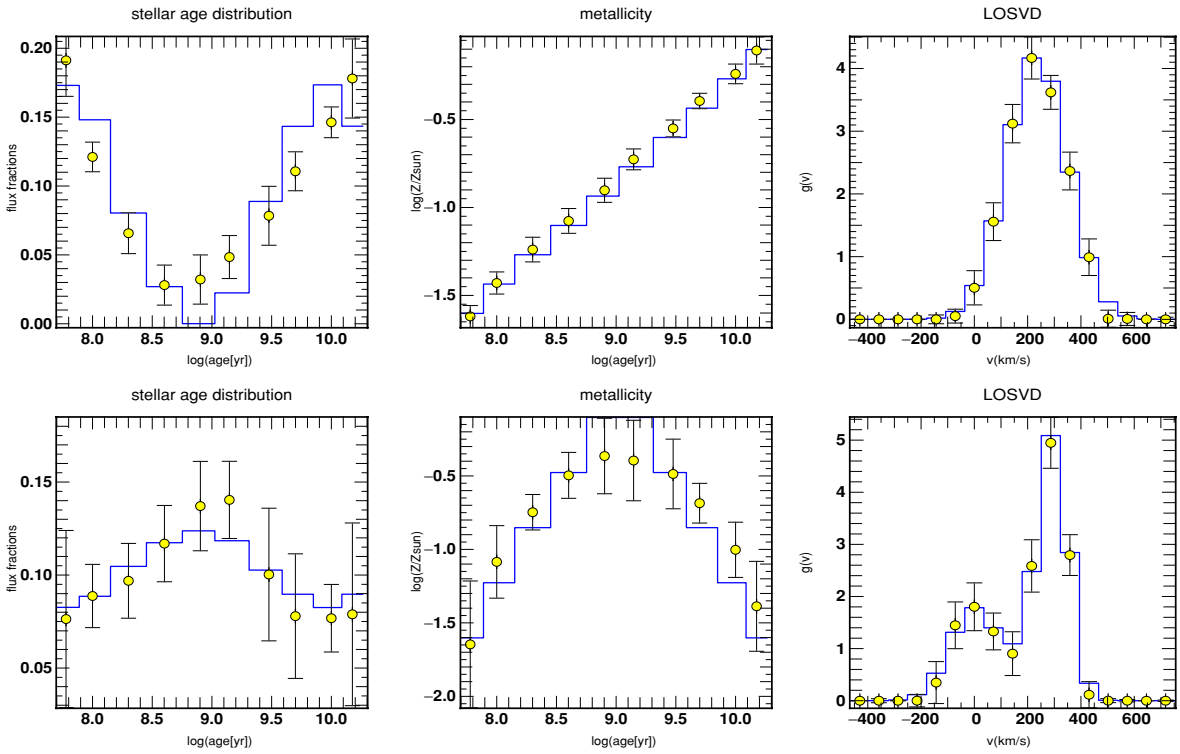

Figure 4. Reconstruction of the stellar age distribution, age–metallicity relation and LOSVD for simulated SDSS-like data with SNR = 30 per pixel. The histogram is the input model. The circles and the bars show respectively the median and the interquartiles of the recovered solutions for 50 realizations.

By inserting the expansion into equation (10) we obtain

$$\phi(w) = \iint \sum_{i=1}^{p} \sum_{j=1}^{n} v_{ij}\theta_{ij}(u,t)B(w-u,t)\,\mathrm{d}t\,\mathrm{d}u,$$

$$\simeq \sum_{i=1}^{p} \sum_{j=1}^{n} v_{ij} B_j(w-u_i). \tag{28}$$

As in the previous sections, $B_j\,(u)$ is a time-averaged SSP of age $t_j \pm \frac{1}{2}\delta t$. We then discretize along wavelengths by averaging over small δw:

$$\phi_k = \frac{1}{\delta w} \sum_{i=1}^{p} \sum_{j=1}^{n} v_{ij} \int B_j(w-u_i)\,\theta\left(\frac{w-w_k}{\delta w}\right)\mathrm{d}w,$$

$$\simeq \sum_{i=1}^{p} \sum_{j=1}^{n} v_{ij} B_j(w_k-u_i), \tag{29}$$

where $(w_j)_{j\in\{0,\ldots,m\}}$ is a set of constant step logarithmic wavelengths. The above expression also reads in matrix form as a sum of kernel convolutions. Finally, the model SED of the emitted light reads

$$\boldsymbol{s} = \sum_{j=1}^{n} \mathbf{K}_j \cdot \boldsymbol{v}_j, \tag{30}$$

where $\boldsymbol{s} = (\phi_1, \phi_2, \ldots, \phi_m)$, $\boldsymbol{v}_j = (v_{1j}, v_{2j}, \ldots, v_{pj})$ and

$$\mathbf{K}_j = \begin{bmatrix} K_{11j} & K_{12j} & \ldots & K_{1pj} \\ K_{21j} & K_{22j} & \ldots & K_{2pj} \\ \vdots & \vdots & \ddots & \vdots \\ K_{m1j} & K_{m2j} & \ldots & K_{mpj} \end{bmatrix}, \tag{31}$$

with

$$K_{ikj} \equiv B_j(w_k - u_i). \tag{32}$$

With this notation, $\mathbf{K}_j$ and $\boldsymbol{v}_j$ are respectively the convolution kernel and the LOSVD of the subpopulation of age t_j, and the model spectrum $\boldsymbol{y}$ is the sum of the convolution of the kernel of each subpopulation with its own LOSVD.

5.2 2D age–velocity smoothness constraints

In the previous sections, the unknowns were 1D functions of time or velocity. Here, the unknown is a 2D distribution, and we thus have to implement a 2D smoothing constraint. We wish to allow the smoothness in age to be distinct from the smoothness in velocity. We thus construct two penalizing functions, P_a and P_v, relying on the standard function P. P_a computes the sum of the Laplacians of the columns of $\mathbf{v}$ while P_v computes the sum of the Laplacians of the lines of $\mathbf{v}$. The smoothness in the direction of the velocities (respectively ages) is set by μ_v (respectively μ_a). We define the vectors $\boldsymbol{v}_j = (v_{1j}, v_{2j}, \ldots, v_{pj})$ as the columns of $\mathbf{v}$, i.e. the LOSVDs of the subpopulations. We similarly define the vectors $\boldsymbol{v}^i = (v_{i1}, v_{i2}, \ldots, v_{in})$ as the lines of $\mathbf{v}$. With this notation, the penalization P_μ reads

$$P_{\boldsymbol{\mu}}(\mathbf{v}) \equiv \mu_a P_a(\mathbf{v}) + \mu_v P_v(\mathbf{v}),$$

$$\equiv \mu_a \sum_{i=1}^{p} P(\boldsymbol{v}^i) + \mu_v \sum_{j=1}^{n} P(\boldsymbol{v}_j). \tag{33}$$

The objective function, Q_μ, is now fully specified as $Q_\mu = \chi^2 + P_\mu$. Its gradients are given in Appendix Section A3. Here we choose the smoothing parameters, $\boldsymbol{\mu} \equiv (\mu_a, \mu_v)$, on the basis of simulations.

5.3 Simulations of a bulge–disc system

We studied the feasibility of separating two age-dynamically distinct populations, i.e. two components that do not overlap in an

Table 1. Projected kinematic parameters and age of the model bulge–disc system used to produce the simulations of Figs 5 and 6. V_c (respectively σ_v) is the rotation velocity (respectively, the velocity dispersion) projected on the line of sight.

	V_c (km s^{-1})	σ_v (km s^{-1})	Age (Gyr)
Case 1			
Bulge	0	100	8
Disc	120	30	0.5
Case 2			
Bulge	0	150	8
Disc	0	50	0.5

age–velocity distribution diagram, in a regime of very high-quality model and data. We performed simulations in the idealized case of a very simplified spiral galaxy consisting of a bulge–disc system of solar metallicity seen without extinction at some intermediate inclination, in two observational contexts. The corresponding ages and projected kinematic parameters are given in Table 1. The resolution of the pseudo-data is $R = 10\,000$ at 4000–6800 Å, and the SNR is 100 per 0.2 Å pixel.

Case 1: The galaxy is resolved, and the fibre aperture is small compared to the angular size of the galaxy. The line of sight is offset by a couple of kiloparsecs from the centre along the major axis. The projected model age–velocity distribution involves two superimposed components: an old, non-rotating, kinematically hot population representing the bulge; and a young, rotating, kinematically cold component. The model and the median of 30 reconstructions are shown in Fig. 5. The separation of the components is clear and their parameters can be recovered with good accuracy, considering the difficulty of the task.

Case 2: The galaxy is unresolved. The difference from the former situation is that, because of the spatial integration, both age–velocity distributions are centred. For a given dynamical model, the projected dispersion of the disc component depends on its inclination. Fig. 6 shows that the separation is successful and that the ages and integrated kinematic properties of both components can be measured.

6 CONCLUSIONS AND OUTLOOK

6.1 Conclusions

The non-parametric kinematic deconvolution of a galaxy spectrum is efficiently performed using a MAP formalism (Section 3). Regularization through smoothness requirements and positivity significantly improve the behaviour of the inversion with respect to noise in the data. This improvement occurs at the cost of introducing some bias in the reconstructed LOSVD, but this bias remains reasonable. Strong non-Gaussianities of LOSVDs are reliably detected from mock data generated using PÉGASE-HR SSPs for SNR down to 20 per 0.2 Å pixel.

When the template does not exactly match the model spectrum at rest, i.e. there is some template mismatch, the error on the velocity dispersion increases very quickly (Section 3.4). For example, in our experiments, where $\sigma_v = 50$ km s^{-1} with $R = 10\,000$ data, the error on the measured velocity dispersion amounts to 10–20 per cent if the template differs from the model by more than 0.3 dex in age and metallicity, perpendicular to the age–metallicity degeneracy.

The formal similarity between the non-parametric kinematic deconvolution and the recovery of the stellar content allows us to merge both processes in a 'mixed' inversion where the observed spectrum is fitted by determining the stellar content and the kinematics simultaneously (Section 4). This circumvents the need for iterations where kinematic and stellar content analyses are carried out one after the other, until convergence is reached; this provides an efficient method to analyse large sets of data.

Satisfactory reconstructions of the stellar age distribution, the age–metallicity relation, the extinction and the global LOSVD were obtained from mock data down to $R = 2000$, SNR = 30 per 1 Å pixel in the 4000–6800 Å range (simulating SDSS data in the PÉGASE-HR range), indicating the good behaviour of the method. Since, in our simulations, the introduction of the kinematics into STECMAP did not affect the recovery of the stellar content, we consider that the

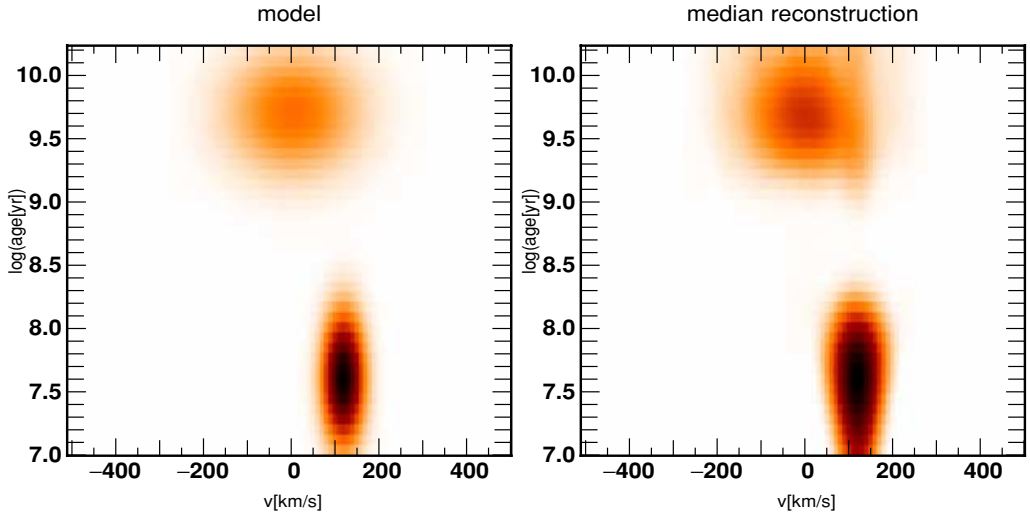

Figure 5. Model (left) and median reconstruction (right, ≈30 realizations) of a stellar age–velocity distribution from SNR = 100 per 0.2 Å pixel pseudo-data at 4000–6800 Å. The model stellar age–velocity distribution mimics that of a simplified spiral galaxy seen with intermediate inclination. The old, broad component can account for the bulge population, while the young, narrow, rotating component represents a thin disc population. The projected kinematic parameters of the model are given in Table 1 (Case 1). The different kinematic components are well separated and clearly identifiable.

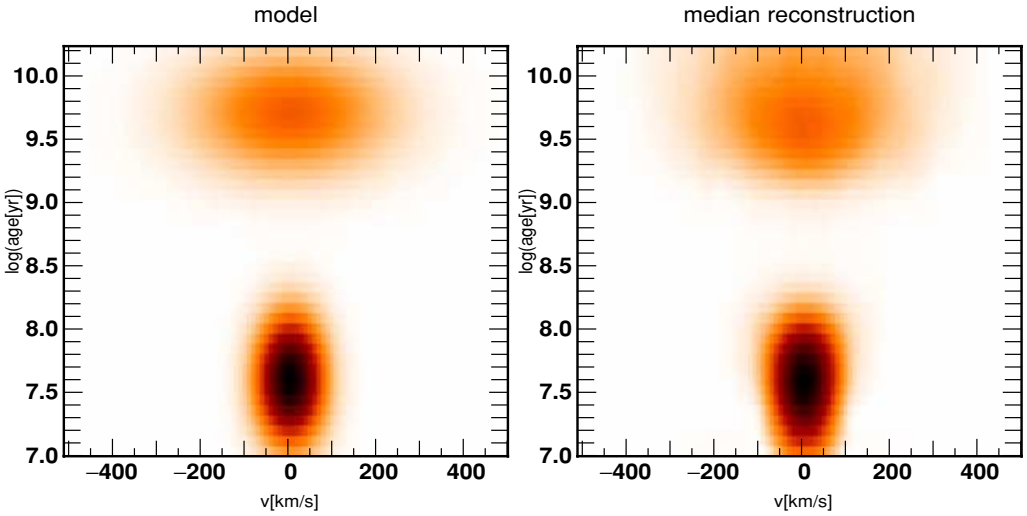

Figure 6. Same as Fig. 5 but for an unresolved simplified spiral galaxy with projected and spatially integrated kinematic parameters given in Table 1 (Case 2). The velocity dispersion of the integrated young disc component depends on the inclination angle. The bulge and the disc are well separated and clearly identifiable. Their respective velocity dispersions and ages are reliably recovered.

error estimates and separability analysis given in Paper I remain valid.

In a more exploratory part of this work, we showed the feasibility of recovering age-dependent kinematics in a simplified monometallic unreddened context (Section 5). We were able to separate the bulge and disc components of a simplified model spiral galaxy in integrated light provided very high-quality data (SNR = 100 per 0.2 Å pixel in the optical domain) and models are available, i.e. we constrain both components in velocity dispersion and age. This separation was also carried out successfully in the setup corresponding to an unresolved galaxy.

Further investigations are needed to extend this technique to a regime where the metallicity and extinction are unknown. We expect that letting the metallicity be a free parameter would certainly lead to a more degenerate problem, as shown by the degradation of the resolution in age found in Paper I compared to fixed metallicity problems. On the contrary, we do not expect the addition of the extinction as a free parameter or a more complex form of extinction law or flux calibration correction, possibly non-parametric, to deteriorate the conditioning of the problem. The results are encouraging, and the feasibility of such age-dependent kinematics reconstructions deserves to be tackled in realistic specific pseudo-observational regimes in the future.

As mentioned in Paper I, the SSP models were considered to be perfect and noiseless. It still has to be investigated how instrumental error sources such as flux and wavelength calibration error, additive noise, contamination by adjacent objects and, equally important, model errors can affect the robustness of such sophisticated interpretations.

6.2 Outlook

STECKMAP will be very useful to interpret data of large spectroscopic surveys, complete or in progress, such as 2DFGRS,[1] SDSS,[2] DEEP2,[3] or VVDS,[4] especially where constraints on both the stellar content and the dynamics are required. STECKMAP' s analysis of the spectroscopic survey data or of an SNR-selected subsample, combined with survey photometry, could provide estimates of the stellar and dynamical masses (which must be corrected for fibre aperture though), thereby allowing astronomers the prospect of investigating the dark matter content in galaxies on a statistically significant sample, in the spirit of Padmanabhan et al. (2004).

The application of age-dependent kinematics to integral field spectroscopy data from, for example, SAURON (Bacon et al. 2001; de Zeeuw et al. 2002), OASIS (McDermid et al. 2004a), MUSE (Henault et al. 2003) or MPFS (Chilingarian et al. 2004) could significantly boost the amount of information extracted from these data.

The inner parts of elliptical or dwarf elliptical galaxies have shown via adaptive optics new kinematically decoupled structures (cores or central discs), which were previously unresolved (McDermid et al. 2004b; Bacon et al. 2001). Similarly, if decoupled structures are unresolved and remain so, even with adaptive optics, it may still be possible to separate components in age–velocity space. Hence, the technique presented in Section 5 extends the range of investigation for the inner components of galaxies even further in redshift and distance with the current generation of instruments. The faint, generalized counterparts of kinematically decoupled cores, i.e. stellar streams generated by minor merging and accretion of satellites, may also be detectable by an age-dependent kinematics reconstruction in systems that cannot be resolved into stars, provided that they are sufficiently distinct from the bulk stars of the galaxy in the age–velocity space. This will enlarge the sample of galaxies for which such detailed information is available, and may make it statistically significant.

ACKNOWLEDGMENTS

We are grateful to A. Siebert for useful comments and helpful suggestions. We would like to thank D. Munro for freely distributing

[1] http://www.mso.anu.edu.au/2dFGRS/
[2] http://www.sdss.org/
[3] http://www.deep.berkeley.edu/
[4] http://www.oamp.fr/virmos/vvds.htm

his Yorick programming language,[5] together with its MPI interface, which we used to implement our algorithm in parallel. PO thanks the MPA for their hospitality and funding from a Marie Curie studentship.

REFERENCES

Bacon R. et al., 2001, MNRAS, 326, 23
Bacon R., Emsellem E., Combes F., Copin Y., Monnet G., Martin P., 2001, A&A, 371, 409
Balcells M., Quinn P. J., 1990, ApJ, 361, 381
Bender R., 1990, A&A, 229, 441
Bender R., Surma P., 1992, A&A, 258, 250
Calzetti D., 2001, PASP, 113, 1449
Cappellari M., Emsellem E., 2004, PASP, 116, 138
Chilingarian I., Prugniel P., Silchenko O., Afanasiev A., 2004, astro-ph/0412293
Davies R. L., Sadler E. M., Peletier R. F., 1993, MNRAS, 262, 650
De Bruyne V., De Rijcke S., Dejonghe H., Zeilinger W. W., 2004a, MNRAS, 349, 440
De Bruyne V., De Rijcke S., Dejonghe H., Zeilinger W. W., 2004b, MNRAS, 349, 461
De Rijcke S., Dejonghe H., Zeilinger W. W., Hau G. K. T., 2004, A&A, 426, 53
de Zeeuw P. T. et al., 2002, MNRAS, 329, 513
Faber S. M., Friel E. D., Burstein D., Gaskell C. M., 1985, ApJS, 57, 711
Falcón-Barroso J., Balcells M., Peletier R. F., Vazdekis A., 2003, A&A, 405, 455
Ferguson A. M. N., Irwin M. J., Ibata R. A., Lewis G. F., Tanvir N. R., 2002, AJ, 124, 1452
Freeman K., Bland-Hawthorn J., 2002, ARA&A, 40, 487
Henault F. et al., 2003, in Iye M., Moorwood A. F. M., eds, Proc. SPIE 4841, Instrument Design and Performance for Optical/Infrared Ground-Based Telescopes. SPIE, Bellingham, WA, p. 1096
Ibata R., Chapman S., Ferguson A. M. N., Irwin M., Lewis G., McConnachie A., 2004, MNRAS, 351, 117
Kroupa P., Tout C. A., Gilmore G., 1993, MNRAS, 262, 545
Kuijken K., Merrifield M. R., 1993, MNRAS, 264, 712
Kuntschner H., 2000, MNRAS, 315, 184
Kuntschner H., 2004, A&A, 426, 737
Le Borgne D., Rocca-Volmerange B., Prugniel P., Lançon A., Fioc M., Soubiran C., 2004, A&A, 425, 881
McDermid R. et al., 2004a, Newsl. Isaac Newton Group Telescopes, No 8, p. 3
McDermid R. et al., 2004b, Astron. Nachr., 325, 100
Merritt D., 1997, AJ, 114, 228
Ocvirk P., Pichon C., Lançon A., Thiébaut E., 2005, MNRAS, in press (Paper I, this volume)
Padmanabhan N. et al., 2004, New Astron., 9, 329
Pichon C., Siebert A., Bienaymé O., 2002, MNRAS, 329, 181
Pinkney J. et al., 2003, ApJ, 596, 903
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 2002, Numerical Recipes in C++: The Art of Scientific Computing, 2nd edn. Cambridge Univ. Press, Cambridge
Rix H., White S. D. M., 1992, MNRAS, 254, 389
Saha P., Williams T. B., 1994, AJ, 107, 1295
Trager S. C., Worthey G., Faber S. M., Burstein D., Gonzalez J. J., 1998, ApJS, 116, 1
van der Marel R. P., Franx M., 1993, ApJ, 407, 525
Wahba G., ed., 1990, Spline Models for Observational Data (CBMS-NSF Regional Conf. Ser. Appl. Math., Vol. 59). SIAM, Philadelphia
Worthey G., 1994, ApJS, 95, 107

[5] Available at http://www.maumae.net/yorick/doc/index.html

APPENDIX A: GRADIENT COMPUTATIONS

A1 Kinematic deconvolution

In this section we derive the gradient of Q_μ with respect to the LOSVD $\boldsymbol{g}$. First, we rewrite the χ^2 term as

$$\chi^2 = \boldsymbol{r}^{\mathrm{T}} \cdot \mathbf{W} \cdot \boldsymbol{r}, \tag{A1}$$

where the residuals vector $\boldsymbol{r}$ is defined by

$$\boldsymbol{r} = \boldsymbol{y} - \mathcal{F}^{-1} \cdot \mathrm{diag}(\mathcal{F} \cdot \boldsymbol{F}) \cdot \mathcal{F} \cdot \boldsymbol{g}. \tag{A2}$$

The derivative of the χ^2 then reads

$$\frac{\partial \chi^2}{\partial \boldsymbol{g}} = -2\mathcal{F}^* \cdot \mathrm{diag}(\mathcal{F} \cdot \boldsymbol{F})^* \cdot \mathcal{F} \cdot \mathbf{W} \cdot \boldsymbol{r}, \tag{A3}$$

where the asterisk * denotes the complex conjugate. Since the stellar template and the LOSVD can play symmetrical roles in equation (17), we can also write the derivative of χ^2 relatively to the stellar template:

$$\frac{\partial \chi^2}{\partial \boldsymbol{F}} = -2\mathcal{F}^* \cdot \mathrm{diag}(\mathcal{F} \cdot \boldsymbol{g})^* \cdot \mathcal{F} \cdot \mathbf{W} \cdot \boldsymbol{r}. \tag{A4}$$

This expression will be useful for later derivations of gradients for more complex problems in the following appendices.

A2 Gradients of the mixed inversion

Here we show how to obtain the partial derivatives of $Q_\mu = \chi^2 + P_\mu$ as defined in Section 4. Given that writing the derivatives of the penalizing functions P_μ is straightforward, in this appendix we will focus on the gradients of χ^2. In the mixed inversion, the reddened model spectrum at rest plays the role of the stellar template $\boldsymbol{F}$ in the classical kinematic deconvolution of equation (15). Thus $\partial\chi^2/\partial\boldsymbol{g}$ can be obtained by replacing $\boldsymbol{F} \leftarrow \mathrm{diag}(\boldsymbol{f}_{\mathrm{ext}}(E)) \cdot \mathbf{B} \cdot \boldsymbol{x}$ in equation (A3):

$$\frac{\partial \chi^2}{\partial \boldsymbol{g}} = -2\mathcal{F}^* \cdot \mathrm{diag}(\mathcal{F} \cdot \mathrm{diag}(\boldsymbol{f}_{\mathrm{ext}}(E)) \cdot \mathbf{B} \cdot \boldsymbol{x})^* \cdot \mathcal{F} \cdot \mathbf{W} \cdot \boldsymbol{r}, \tag{A5}$$

where $\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{s}$ is the residuals vector, with $\boldsymbol{s}$ as given by equation (25). To obtain the other partial derivatives, we use the following relation. For any parameter α we have

$$\frac{\partial \chi^2}{\partial \alpha} = \left(\frac{\partial \chi^2}{\partial \boldsymbol{F}}\right)^{\mathrm{T}} \cdot \frac{\partial \boldsymbol{F}}{\partial \alpha}. \tag{A6}$$

The first term $\partial\chi^2/\partial\boldsymbol{F}$ is given by equation (A4), while the second term reads, considering each unknown,

$$\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}} = \mathrm{diag}(\boldsymbol{f}_{\mathrm{ext}}) \cdot \mathbf{B}, \tag{A7}$$

$$\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{Z}} = \mathrm{diag}(\boldsymbol{x}) \cdot \frac{\partial \mathbf{B}}{\partial \boldsymbol{Z}} \cdot \mathrm{diag}(\boldsymbol{f}_{\mathrm{ext}}), \tag{A8}$$

$$\frac{\partial \boldsymbol{F}}{\partial E} = \mathrm{diag}\left(\frac{\partial \boldsymbol{f}_{\mathrm{ext}}}{\partial E}\right) \cdot \mathbf{B} \cdot \boldsymbol{x}, \tag{A9}$$

with the same notation as in the appendix of the STECMAP paper.

A3 Gradients for the age-dependent kinematics recovery

Again, we focus on the partial derivatives of χ^2. Using equation (17), the model can be rewritten using the Fourier operator

$$\boldsymbol{s} = \sum_{j=1}^{n} \mathcal{F}^* \cdot \mathrm{diag}(\mathcal{F} \cdot \boldsymbol{B}_j) \cdot \mathcal{F} \cdot \boldsymbol{v}_j, \tag{A10}$$

where $\boldsymbol{B}_j$ is the discretized time-averaged SSP of age $[t_{j-1}, t_j]$. The derivatives of χ^2 relative to $\mathbf{v}$ can be derived directly from equation (A3) since the model is just a sum of convolutions. Replacing $\boldsymbol{F} \leftarrow \boldsymbol{B}_j$ and $\boldsymbol{g} \rightarrow \boldsymbol{v}_j$ yields the gradient of χ^2:

$$\frac{\partial \chi^2}{\partial \boldsymbol{v}_j} = -2\mathcal{F}^* \cdot \mathrm{diag}(\mathcal{F} \cdot \boldsymbol{B}_j)^* \cdot \mathcal{F} \cdot \mathbf{W} \cdot \boldsymbol{r}, \tag{A11}$$

with the residuals vector $\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{s}$. Finally, the derivative of Q_μ relative to $\mathbf{v}$ is the matrix defined by

$$\frac{\partial Q_\mu}{\partial \mathbf{v}} = \left(\frac{\partial Q_\mu}{\partial \boldsymbol{v}_1}, \frac{\partial Q_\mu}{\partial \boldsymbol{v}_2}, \ldots, \frac{\partial Q_\mu}{\partial \boldsymbol{v}_n} \right). \tag{A12}$$

This paper has been typeset from a TEX/LATEX file prepared by the author.