Tracking the Temporal-Evolution of Supernova Bubbles in Numerical Simulations

Marco Canducci[1(✉)], Abolfazl Taghribi[2], Michele Mastropietro[3], Sven de Rijcke[3], Reynier Peletier[2], Kerstin Bunte[2], and Peter Tino[1]

[1] University of Birmingham, Edgbaston B15 2TT, Birmingham, UK
{M.Canducci,p.tino}@bham.ac.uk
[2] Faculty of Science and Engineering, University of Groningen, Groningen, The Netherlands
{a.taghribi,k.bunte}@rug.nl, peletier@astro.rug.nl
[3] Department of Physics and Astronomy, Ghent University, Ghent, Belgium
{michele.mastropietro,sven.derijcke}@ugent.be

Abstract. The study of low-dimensional, noisy manifolds embedded in a higher dimensional space has been extremely useful in many applications, from the chemical analysis of multi-phase flows to simulations of galactic mergers. Building a probabilistic model of the manifolds has helped in describing their essential properties and how they vary in space. However, when the manifold is evolving through time, a joint spatio-temporal modelling is needed, in order to fully comprehend its nature. We propose a first-order Markovian process that propagates the spatial probabilistic model of a manifold at fixed time, to its adjacent temporal stages. The proposed methodology is demonstrated using a particle simulation of an interacting dwarf galaxy to describe the evolution of a cavity generated by a Supernova.

Keywords: Manifold learning · l1 and l2-regularization · Temporal generative topographic mapping · SPH simulation · Superbubble

1 Introduction

Generally, manifold learning assumes relevant information to lie on possibly noisy, low-dimensional structures embedded in higher dimensions. In order to extract this information it is essential to build a robust representation of the structure itself. Many approaches have been proposed, based on different assumptions, that address such a problem. Locally Linear Embedding (LLE) [9], Multi-Dimensional Scaling (MDS) [10] and their generalizations, aim at recovering a low-dimensional representation of the embedded manifold by covering it

Supported from the European Union's Horizon 2020 Marie Sklodowska-Curie grant agreement No. 721463 and the Alan Turing Institute Fellowship n. 96102.

H. Yin et al. (Eds.): IDEAL 2021, LNCS 13113, pp. 493–501, 2021.
https://doi.org/10.1007/978-3-030-91608-4_49

with hyper-planes (such as LLE) or building an adjacency matrix of locally linear neighbourhoods (like MDS). However, if the manifold has inherent transverse noise (e.g. measurement errors) Generative Topographic Mapping (GTM) [1], as a fully probabilistic principled model, is preferred. GTM naturally handles and preserves the information retained by noise while producing a low-dimensional latent representation of the manifold's topological structure.

In several situations the nature of the manifold is dynamic and it evolves through time, e.g. astronomical numerical simulations. In these cases, a joint spatio-temporal representation of the manifold is needed, in order to capture its underlying nature. It has been shown how GTM can be generalized to model time-series (GTM Through Time) [4], by imposing a Hidden Markov Model (HMM) structure on latent space nodes [8]. This enables one to project whole trajectories evolving around a lower-dimensional manifold onto the latent space of a single GTM. Here we tackle a different problem - we propose a methodology for tracking the evolution of a time series of low-dimensional manifolds based on the first-order Markovian assumption [5], where the spatial structure of the manifold at each time-step is modelled via GTM. The model is propagated through adjacent time-steps by minimization of a set of cost functions, aimed at preserving the position of the Gaussian Mixture's centres, while still fitting the evolved data set. The proposed regularizations are based on L_1- and L_2-norms on changes of model parameters. The methodology is applied to a particle simulation of a dwarf galaxy interacting with its host galaxy cluster. The manifold of interest is a cavity ("bubble") generated by a Supernova and detected via the Dionysus toolbox [7] (Persistent Homology, PH) at a given temporal stage. In [3] a generalization of GTM to abstract graphs has been proposed. However, knowing the topological properties of cavities, we adopt Geodesic GTM [11] as a modelling technique.

2 Methodology

Consider a sequence of $\mathcal{S} = (\mathcal{Q}^0, \ldots, \mathcal{Q}^\tau, \mathcal{Q}^{\tau+1}, \ldots, \mathcal{Q}^T)$ of point clouds $\mathcal{Q}^\tau = \left\{\mathbf{t}_1^\tau, \mathbf{t}_2^\tau, \ldots, \mathbf{t}_{N_\tau}^\tau\right\} \subset \mathbb{R}^d$, sampling a noisy, time-dependent, low-dimensional manifold $\mathcal{M}(\tau)$ of dimension ℓ throughout its evolution, $\tau = 1, 2, ..., T$. A naive approach to produce a probabilistic model of the sequence $\mathcal{S}$ is to model each temporal realization of the manifold individually, disregarding the mutual temporal dependencies of its components. However, it is sometimes useful to include the time-dependency of data sets when interested in joint space-time variations of the manifold's properties. We first describe the chosen methodology for modelling the spatial state of manifold $\mathcal{M}(\tau)$ for two topologically different cases. We then propose different measures for the propagation of such models through adjacent temporal stages.

2.1 One-Dimensional GTM

Let us assume that the manifolds $\mathcal{M}(\tau)$ are open, one-dimensional and nonlinear. As previously mentioned, each $\mathcal{M}(\tau)$ is sampled by $\mathcal{Q}^\tau$. Since we have the

information about the manifold's dimensionality ($\ell = 1$), we take advantage of the Generative Topographic Mapping (GTM) formulation and define our model as a constrained mixture of Gaussians, where the latent space representation of centres is the interval $\mathrm{I} = [-1, 1]$. We sample I regularly with M centres $\mathbf{X} = \{x_1, \ldots, x_M\}$ and define a set of $K < M$ Radial Basis Functions (RBFs) $\phi(\mu_k, x_m)$ centred on uniformly spaced points $\mathbf{K} = \{\mu_1, \ldots, \mu_K\} \in \mathrm{I}$. We choose the RBFs to be spherical Gaussians:

$$\phi(\mu_k, x_m) = \exp\left[-\frac{(\mu_k - x_m)^2}{2\sigma^2}\right] \ , \tag{1}$$

where σ is a scale parameter chosen to be a multiple of the distance between neighboring RBFs centres. We then map the set of latent centres $\mathbf{X}$ onto the ambient space of manifold $\mathcal{M}(\tau)|_{\tau=0}$ through the non-linear mapping function $\mathbf{y}(\mathbf{X}; \mathbf{W})$ defined as:

$$\mathbf{y}(\mathbf{X}; \mathbf{W}) = \mathbf{W}\Phi(\mathbf{X}, \mathbf{K}) \ . \tag{2}$$

Here, Φ is the matrix with $\phi(\mu_k, x_m)$ being its $(k, m)-$th element. Each $x_m \in \mathbf{X}$ is thus mapped, through Eq. (2), to point $\mathbf{c}_m \in \mathbb{R}^d$, which acts as a centre of the Gaussian mixture. On each mapped centre $\mathbf{c}_m$ we impose a manifold-aligned noise model obtained by differentiation of the mapping function $\mathbf{y}(\mathbf{X}; \mathbf{W})$ w.r.t. the latent space coordinates, as described in [2]:

$$\mathbf{C}_m = \frac{1}{\beta}\mathbf{I} + \eta_m \sum_{p=1}^{\ell} \left[\left(\frac{\partial \mathbf{y}}{\partial x_p}\right)^{\top} \left(\frac{\partial \mathbf{y}}{\partial x_p}\right)\right]\Bigg|_{x_m} \ . \tag{3}$$

Here, x_p denotes the latent coordinate of dimension p. With suitable values for parameters β and η_m $(m = 1, \ldots, M)$ the density of each component of the mixture reads:

$$p(\mathbf{t}|x_m, \mathbf{W}) = \left(\frac{1}{2\pi}\right)^{d/2} |\mathbf{C}_m|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{c}_m - \mathbf{t})^{\top}\mathbf{C}_m^{-1}(\mathbf{c}_m - \mathbf{t})\right] \ . \tag{4}$$

The overall GTM model is a flat mixture of Gaussians,

$$p(\mathbf{t}|\mathbf{W}) = \frac{1}{M}\sum_{m=1}^{M} p(\mathbf{t}|x_m, \mathbf{W}) \ . \tag{5}$$

whose parameters can be trained via a maximum-likelihood optimization, e.g. through the Expectation-Maximization (EM) algorithm [2].

2.2 Geodesic GTM for Cavities

In [11], we proposed a *Geodesic-GTM*, where the linear latent space of classical GTM is replaced by a spherical one with dedicated geodesic distance. Hence, every point $\boldsymbol{\xi}$ on the surface of the sphere is uniquely determined by a pair of angular coordinates: θ and λ where $(\theta, \lambda) \in \mathrm{I}_{\angle}^{\ell} = [-\pi; \pi] \times [-\pi/2; \pi/2]$ and $\mathrm{I}_{\angle}^{\ell}$

denotes the ℓ-dimensional, angular interval ($\ell = 2$). Assuming a radius $r = 1$ the geodesic distance between any pair of points $\boldsymbol{\xi}_i$ and $\boldsymbol{\xi}_j$ can be rewritten:

$$d_{\Omega}(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j) = 2 \arcsin\left(\frac{\Delta\Omega}{2}\right), \qquad \Delta\Omega = \sqrt{\Delta X^2 + \Delta Y^2 + \Delta Z^2}\ , \tag{6}$$

where $\Delta X = \sin\theta_i \cos\lambda_i - \sin\theta_j \cos\lambda_j$, $\Delta Y = \sin\theta_i \sin\lambda_i - \sin\theta_j \sin\lambda_j$ and $\Delta Z = \cos\theta_i - \cos\theta_j$. As in the previous section, we define on the latent space a set $\mathbf{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M\} \subset \mathrm{I}_{\angle}^{\ell}$ of latent points and $\mathbf{K} = \{\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K\} \subset \mathrm{I}_{\angle}^{\ell}$ of centres for the RBFs $\boldsymbol{\phi} = \{\phi(\boldsymbol{\mu}_1), \ldots, \phi(\boldsymbol{\mu}_K)\}$. Using this formulation each RBF takes the form:

$$\phi(\boldsymbol{\mu}_k, \boldsymbol{\xi}_m) = \exp\left[-\frac{d_{\Omega}(\boldsymbol{\mu}_k, \boldsymbol{\xi}_m)^2}{2\sigma^2}\right]\ . \tag{7}$$

The probabilistic model is then formed analogously to the one-dimensional case, with the only exception for the estimation of the covariance matrix of the centres, that is now dependant on the geodesic distance on a sphere (Eq. (6)). The covariance matrix of each centre $\mathbf{y}(\boldsymbol{\xi}_{\boldsymbol{m}}; \mathbf{W})$ of the Gaussian mixture model is obtained via Eq. (3), where $\ell = 2$, $x_1 = \theta$ and $x_2 = \lambda$.

2.3 Temporal Evolution

Given a data set $\mathcal{Q}^{\tau} = \left\{\mathbf{t}_1^{\tau}, \mathbf{t}_2^{\tau}, \ldots, \mathbf{t}_{N_{\tau}}^{\tau}\right\} \subset \mathbb{R}^d$ sampling the $\ell-$dimensional manifold $\mathcal{M}(\tau)$ at time τ, we now wish to model its evolution in time by propagating the model created at time $\tau - 1$ through the sequence $\mathcal{S}(\tau)$. We impose the temporal smoothness of model structure by chaining the model parameters in the first-order Markov dependency structure through regularization terms. Once the probabilistic model of data set $\mathcal{Q}^{\tau-1}$ is obtained, when presented with the next data set $\mathcal{Q}^{\tau}$, we require minimization of a cost function for the recovery of the updated parameters $\mathbf{W}^{\tau}$. Multiple choices could be made for the cost function; in the following we will present and compare four options that behave differently when applied iteratively on each pair of adjacent data sets in $\mathcal{S}$:

$$\mathrm{l}_{2\mathrm{W}}(\mathbf{W}^{\tau}; \Lambda) = -\mathcal{L}(\mathcal{Q}^{\tau}|\mathbf{X}, \mathbf{W}^{\tau}) + \Lambda\|\mathbf{W}^{\tau} - \mathbf{W}^{\tau-1}\|_2^2 \tag{8}$$

$$\mathrm{l}_{2\mathrm{C}}(\mathbf{W}^{\tau}; \Lambda) = -\mathcal{L}(\mathcal{Q}^{\tau}|\mathbf{X}, \mathbf{W}^{\tau}) + \Lambda\|\mathbf{W}^{\tau}\Phi - \mathbf{W}^{\tau-1}\Phi\|_2^2 \tag{9}$$

$$\mathrm{l}_{1\mathrm{W}}(\mathbf{W}^{\tau}; \Lambda) = -\mathcal{L}(\mathcal{Q}^{\tau}|\mathbf{X}, \mathbf{W}^{\tau}) + \Lambda\|\mathbf{W}^{\tau} - \mathbf{W}^{\tau-1}\|_1 \tag{10}$$

$$\mathrm{l}_{1\mathrm{C}}(\mathbf{W}^{\tau}; \Lambda) = -\mathcal{L}(\mathcal{Q}^{\tau}|\mathbf{X}, \mathbf{W}^{\tau}) + \Lambda\|\mathbf{W}^{\tau}\Phi - \mathbf{W}^{\tau-1}\Phi\|_1\ . \tag{11}$$

Here $\mathcal{L}(\mathcal{Q}|\mathbf{X}, \mathbf{W})$ is the log likelihood of the GTM model with latent structure $\mathbf{X}$ and parameters $\mathbf{W}$, given a data set $\mathcal{Q}$. The cost functions are based on the L_1 and L_2-norms ($\|\cdot\|_1$ and $\|\cdot\|_2$) of the (vectorized) differences between weights or centres at times $\tau - 1$ and τ and are minimized following the simplex search method presented in [6]. Each time we evaluate the cost functions we also update the covariance matrices of the components according to Eq. (3), keeping fixed parameters β and η_m for $m = 1, \ldots, M$.

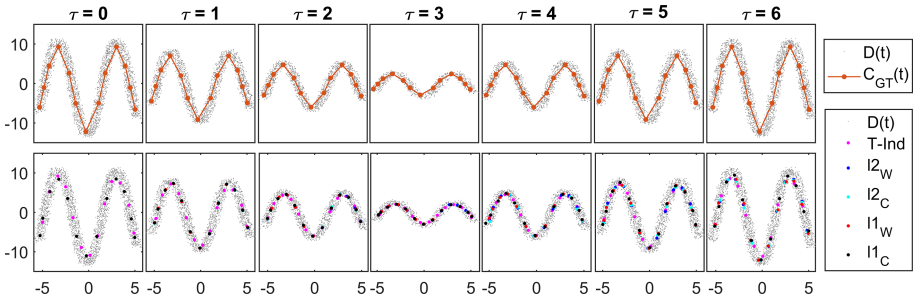

Fig. 1. Top row: 1D time-evolving data set with corresponding "Ground Truth". Bottom: centres obtained via independent (magenta line) or time-dependent modelling.

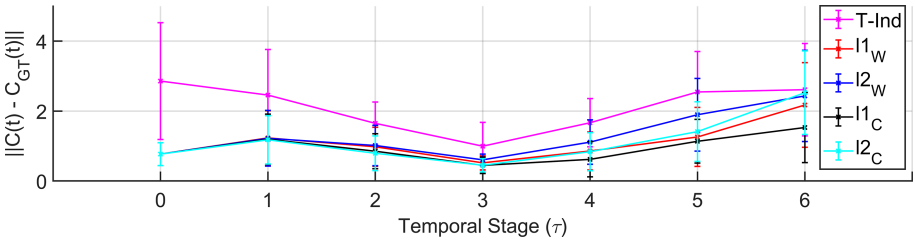

Fig. 2. Mean and standard deviation of distances between updated centres and corresponding Ground Truth at time τ.

3 Experiments on a Synthetic Data Set

In order to test the performance of the minimization procedure when presented with the four cost functions, we constructed a controlled experiment where the Ground Truth is known. A synthetic sequence $\mathcal{S}(\tau)$ is presented in Fig. 1 (top row). Each panel shows the spatial distribution of points at the corresponding time (grey points) and a pre-defined selection of centres considered optimal (Ground Truth: $C_{GT}(\tau)$) for the desription and modelling of the temporal evolotion. For example, to understand the evolution of such "wave shapes", one would need to consistently track the extremal "knot points" and the inner nodes in between them. The one-dimensional GTM presented in Sect. 2.1 is set with 13 centres and 7 RBFs. The weights of the mapping function $\mathbf{y}(\mathbf{X}, \mathbf{W})$ are initialized via linear regression imposing congruence between the mapped centres and $C_{GT}(\tau = 0)$ at stage $\tau = 0$. The Λ hyper-parameter was estimated through 5−fold cross-validation. For comparison, we also modelled each temporal stage individually with the same initialization, but additionally trained with the EM-algorithm. The centres recovered via both time-independent (T-Ind) and dependent modelling are shown in the bottom row of Fig. 1. We also compute at each temporal stage, the distance between the Ground Truth centres at that time and the centres obtained by each modelling approach, with results presented

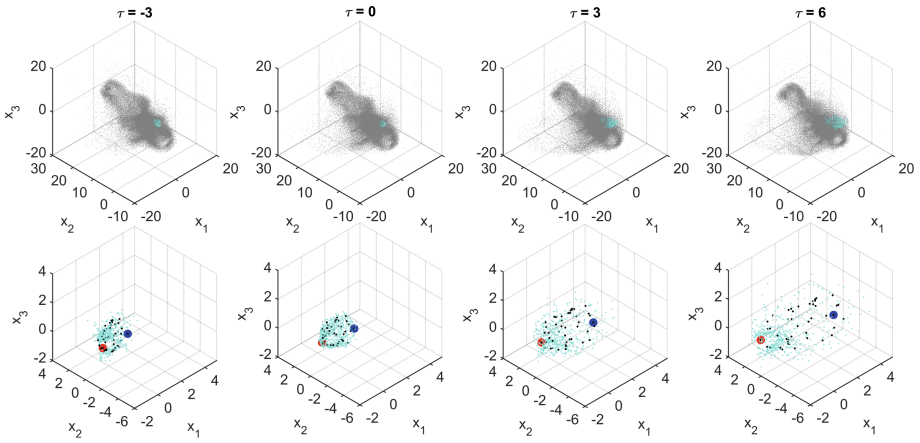

Fig. 3. Top row: Temporal stages of simulated dwarf galaxy. Grey dots are galaxy's point distribution, in cyan the recovered bubble propagated through time. Bottom row: Zoom in on the bubble with centres (black) representing opposite lobes (red and blue). (Color figure online)

in Fig. 2. We show here the mean and standard deviation of this measure at each temporal stage. The minimum distance is consistently achieved by the l1-regularization on centres (l_{1C}), followed in order by l_{1W}, l_{2C}, l_{2W} and finally time-independent modelling. Given the results on this simple synthetic experiment, we infer that $l1-$regularization over the mixture's centres is particularly convenient when identifiability of local properties of a manifold through time is required.

4 Tracking Supernovae Explosions Through Time

We now demonstrate our methodology in the context of a simulated dwarf galaxy, orbiting around the halo of its host galaxy cluster. By increasing local densities the interaction between the gas of the galaxy and that of the cluster may enforce the formation of massive stars which later approach the phase of Supernovae. When a star reaches this stage it expands outwards in a short time, sweeping away the surrounding gas. The "explosion" of the Supernova generates a cavity in the ambient gas, usually referred to as a "bubble". In [7], a methodology for detecting cavities has been proposed using PH. In [11] a sub-sampling procedure (ASAP) has been shown to regularize the detection through PH and applied to a similar astronomical simulation. We aim to track the evolution of a detected bubble through adjacent temporal-stages of the particle simulation. To preserve the position of centres in adjacent temporal stages while capturing the topological properties of the bubble, we need to adopt the probabilistic modelling technique proposed in Sect. 2.2. We adopt all four cost functions tested in the 1D case. However, since the $l1-$regularization on centres proved to be the optimal choice

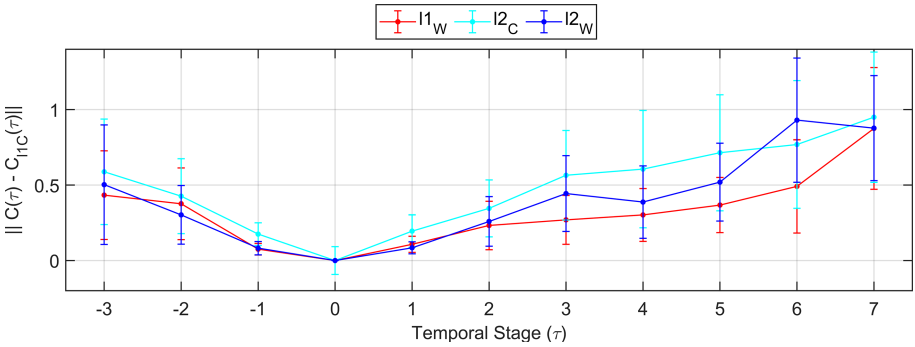

Fig. 4. As in Fig. 2 but w.r.t. the centres obtained via $l1-$regularization "on centres", for the remaining optimization formulations.

in our previous analysis, the detailed comparison will be performed w.r.t. this approach only. We regularly sample the latent space $\mathrm{I}^{\ell}_{\angle}$ with 45 centres and 25 RBFs, initializing the weights of the mapping function by assuming a spherical distribution of the embedded centres[1].

We first initialize the model at time $\tau = 0$ and run the EM-algorithm in order to obtain an accurate model at fixed time of the cavity. We then propagate the model from time $\tau = 0$ to the following and preceding stages, obtaining the spatio-temporal model of the bubble throughout 11 consecutive snapshots. A selection of snapshots equally spaced in time is shown in Fig. 3, top row. Grey dots represent the position of the simulated particles and cyan ones show the bubble's position w.r.t. the galaxy, at each temporal stage. The bubble's particles, detected at $\tau = 0$ via the Dionysus toolbox and ASAP, have been tracked through time using their unique identifiers assigned by the simulation. The bottom row shows a zoom in on the bubble's particles at the corresponding stage with the centres of the respective models in black. We additionally highlight two centres, lying on opposite lobes of the cavity. The red one is closest to the galaxy while the blue one points outwards.

The model's centres have been propagated through adjacent temporal snapshots using the four cost functions defined in Eqs. (8)–(11). Figure 4 shows the mean and standard deviation, at each temporal stage (τ), of the distances from all centres recovered by Eqs. (10) (red), (9) (cyan) and (8) (blue), to the corresponding ones obtained via Eq. (11). Except for a general agreement at temporal stage $\tau = 0$, when the models are initialized, the centres obtained via the three remaining cost functions increasingly diverge with time (both forward and backward). For each selected centre, we can now study how properties of the particles change through time nearby the two selected centres, for all adopted cost functions. In particular, we consider density ρ, temperature T and Neutral Fraction (NF). The Neutral Fraction is the ratio of the neutral Hydrogen mass to the

[1] The radius of the embedded sphere is estimated from the data similar to [11].

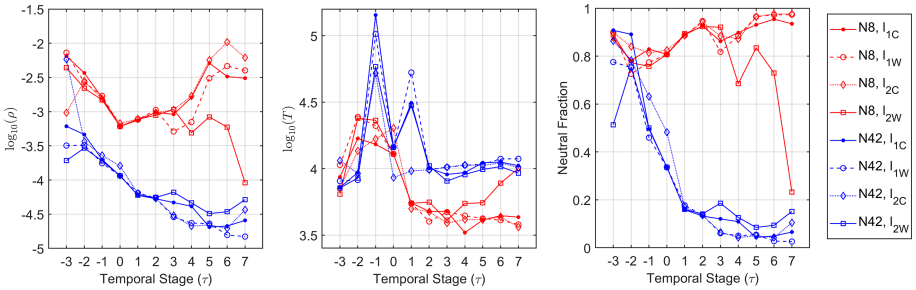

Fig. 5. Weighted mean of simulated quantities ($\log_{10}(\rho)$, T and Neutral Fraction (NF)), on centres $N8$ and $N42$ of the models propagated through time using costs (8)–(11).

total particle mass. For each temporal stage (and each propagated model), we compute the weighted mean of quantity a at centre c_m as:

$$\langle a \rangle|_{\mathbf{c}_m} = \frac{\sum_{n=1}^{N_\tau} a_n R_{mn}}{\sum_{n=1}^{N_\tau} R_{mn}} \quad , \tag{12}$$

where R_{mn} is the *responsibility* of the m-th Gaussian component (around $\mathbf{c}_m$) for the n-th particle and N_τ the number of particles belonging to the bubble at time τ. When applied to properties $a \in \{\rho, T, NF\}$, at each temporal stage for both centres, we obtain the profiles shown in Fig. 5. The results confirm the intuitive expectation of $l1-$regularization to be preferable to the $l2$ one. In addition, direct $l2-$regularization of the centers' position stabilizes the results to a greater degree than $l2-$regularization on $\mathbf{W}$. Centre 8 being closer to the galaxy, has a consistently higher ρ and NF, while retaining lower T w.r.t. to centre 42 that is further away. The cluster's gas is generally hotter and less dense than the gas of the galaxy. The hot gas connected with the blue centre being pushed out can escape the galaxy, keeping a low neutral fraction and high temperature (the UV background is shining on it, keeping it hot and ionized). Gas at the red centre is inside the galaxy's interstellar medium, sheltered from the external UV and can cool and condense to higher densities, which is, indeed, favorable for enhancing star formation.

5 Conclusion

We propose a spatio-temporal probabilistic modelling of time-varying, noisy low dimensional manifolds based on the Generative Topographic Mapping (GTM). The centres of the constrained Gaussian Mixture are propagated through time by minimization of the regularized negative log-likelihood. We study the effect of four different regularizations on the model's parameters, finding that l1-regularization on centres yields the best results in a synthetic experiment. We then demonstrate our methodology to the astronomical particle simulation of a dwarf galaxy interacting with its host galaxy cluster, focusing on a Supernova

bubble. We study the evolution of relevant quantities around selected centres of the Mixture and we find Supernova feedback may be enhancing Star Formation in the galaxy's interior.

References

1. Bishop, C.M., Svensén, M., Williams, C.K.I.: GTM: the generative topographic mapping. Neural Comput. **10**, 215–234 (1998)
2. Bishop, C.M., Svensén, M., Williams, C.K.I.: Developments of the generative topographic mapping. Neurocomputing **21**(1), 203–224 (1998). https://doi.org/10.1016/S0925-2312(98)00043-5
3. Canducci, M., Tiño, P., Mastropietro, M.: Probabilistic modelling of general noisy multi-manifold data sets. Artif. Intell. **302**, 103579 (2022). https://doi.org/10.1016/j.artint.2021.103579, https://www.sciencedirect.com/science/article/pii/S0004370221001302
4. C.M. Bishop, G.E. Hinton, I.S.: GTM through time. In: IET Conference Proceedings, pp. 111–116(5) (1997)
5. Gagniuc, P.A.: Markov Chains: From Theory to Implementation and Experimentation. Wiley, Hoboken (2017). https://doi.org/10.1002/9781119387596
6. Lagarias, J.C., Reeds, J.A., Wright, M.H., Wright, P.E.: Convergence properties of the nelder-mead simplex method in low dimensions. SIAM J. Optim. **9**(1), 112–147 (1998). https://doi.org/10.1137/S1052623496303470
7. Morozov, D.: Dionysus, a C++ library for computing persistent homology. https://mrzv.org/software/dionysus/
8. Rabiner, L.: A tutorial on hidden markov models and selected applications in speech recognition. Proc. IEEE **77**(2), 257–286 (1989). https://doi.org/10.1109/5.18626
9. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. Science **290**, 2323–2326 (2000)
10. Shepard, R.N.: The analysis of proximities: multidimensional scaling with an unknown distance function. i. Psychometrika **27**(2), 125–140 (1962). https://doi.org/10.1007/BF02289630
11. Taghribi, A., Canducci, M., Mastropietro, M., De Rijcke, S., Bunte, K., Tino, P.: ASAP - a sub-sampling approach for preserving topological structures modeled with geodesic topographic mapping. Neurocomputing (2021). https://doi.org/10.1016/j.neucom.2021.05.108