Search for dark substructure in Gaia stellar field using CNN

Mihael Petač

Astroparticle physics group
SISSA, Trieste

April 10, 2019

Motivation

Cold Dark Matter works very well on large scales

- Stable & dark
- Non-relativistic
- Collisionless

DM particle properties encoded in small scales:

- Self-interactions
- Thermal velocities
- Light fermions / ultra-light bosons

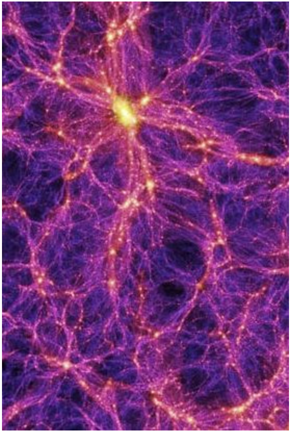


Figure: Millenium simulation

Motivation

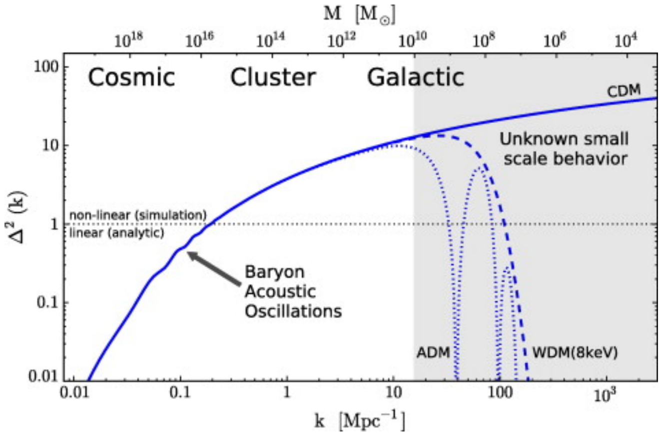


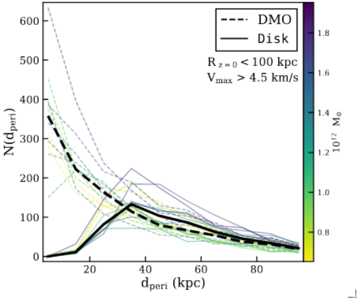
Figure: Kuhlen et al., arXiv:1209.5745

DM Subhalos in Milky Way

Large uncertainties in galactic subhalo abundance:

- DM properties
- Tidal disruption
- Baryonic effects

T. Kelly et al., arXiv:1811.12413



Bosch & Ogiya: MNRAS(2018)475,3 / arXiv:1801.05427

Simulations limited by finite resolution!

Peñarrubia & et al.: MNRAS(2010)406,2 / arXiv:1002.3376

Remnants of cuspy halos could survive

Galactic stellar field

Perturbations of galactic stellar field due to substructure

Gaps in tidal steams, Erkal & Belokurov: MNRAS(2015)454,4 / arXiv:1507.05625

Stellar Wakes from DM subhalos, Buschmann & et al.: PRL(2018)120,211101 / arXiv:1711.03554

Stochastic tidal heating, Peñarrubia & et al.: MNRAS(2019)484,4 / arXiv:1901.11536

Gaia mission DR2:

- Positions and proper motions of $\mathcal{O}(10^7)$ stars
- Accuracy: $\delta v_{\rm los} \sim 1$ km/s

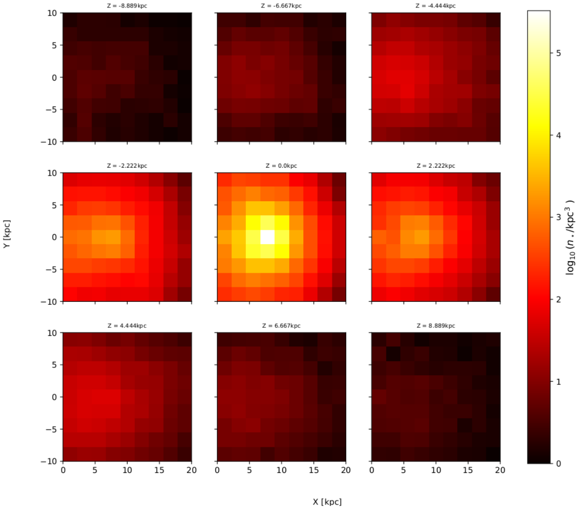
$$\delta v_{\rm t} \sim 1\ \mu\text{arcs/yr} \approx 2\ \text{km/s} \cdot \left(\frac{D_\star}{\rm kpc}\right)$$

Typical size perturbation:

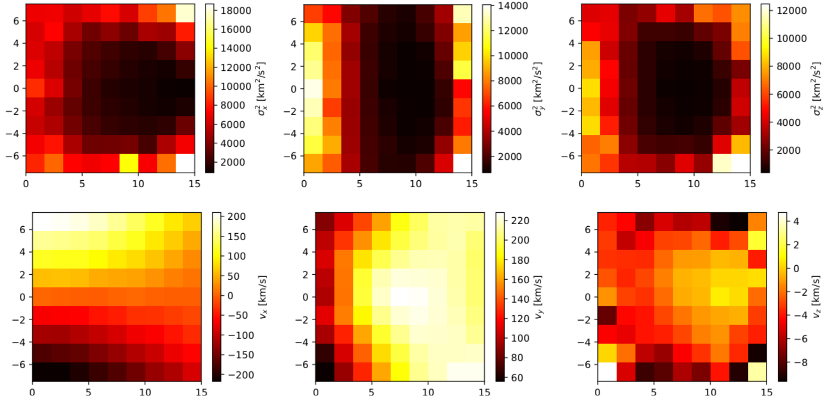
$$\delta v^2 \sim \frac{GM}{\Delta r} \approx 43\ \text{km/s} \cdot \left(\frac{M}{10^7 M_\odot}\right)\left(\frac{\rm kpc}{\Delta r}\right), \quad \Delta r^{-1} = r_1^{-1} - r_0^{-1}$$

Note: $\mathrm{Var}(n_i) = n(\vec{r}_i)\ ,\ \mathrm{Var}(\bar{v}_i) = \frac{\sigma^2(\vec{r}_i)}{n(\vec{r}_i)}\ ,\ \mathrm{Var}(\sigma_i^2) = \frac{2\sigma^4(\vec{r}_i)}{n(\vec{r}_i)-1}$

Galactic stellar field



Galactic stellar field



Signatures in stellar field

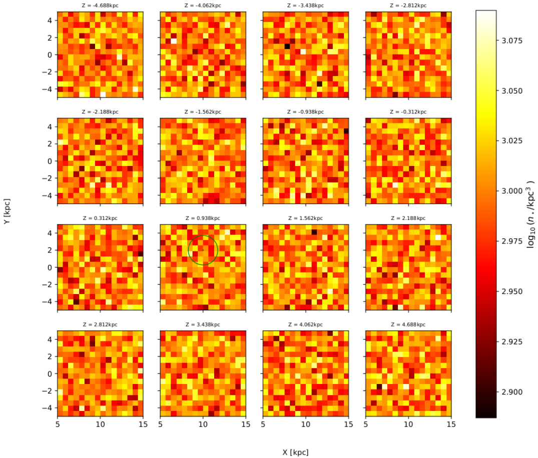
Homogeneous and isotropic stellar field

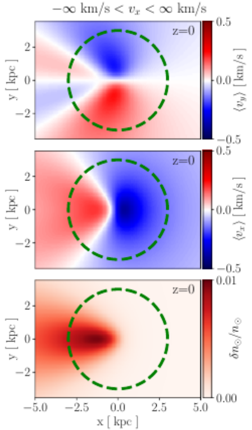
1. Assume: $\bar{f}_\star(\vec{v}) = \frac{n_0}{\sqrt{(2\pi\sigma_\star^2)^3}} \exp\left(-\frac{(\vec{v}-\vec{v}_{\rm sub})^2}{2\sigma_\star^2}\right)$
2. Solve stationary Boltzmann equation for:
   $f_\star(\vec{r},\vec{v}) = \bar{f}_\star(\vec{v}) + \delta f_\star(\vec{r},\vec{v})$, given $\Phi_{\rm sub}(r)$

Realistic stellar field - "painting-in" DM substructure

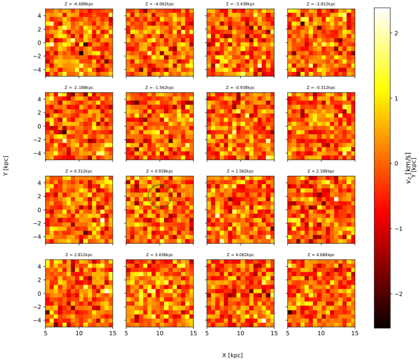
1. Interpolate: $n(\vec{r})$, $\bar{v}(\vec{r})$, $\sigma^2(\vec{r})$
2. Generate stochastic stellar field
3. Modify trajectories – scattering in central potential $\Phi_{\rm sub}(\vec{r})$

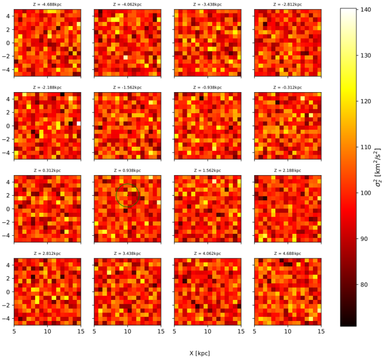
Signatures in stellar field



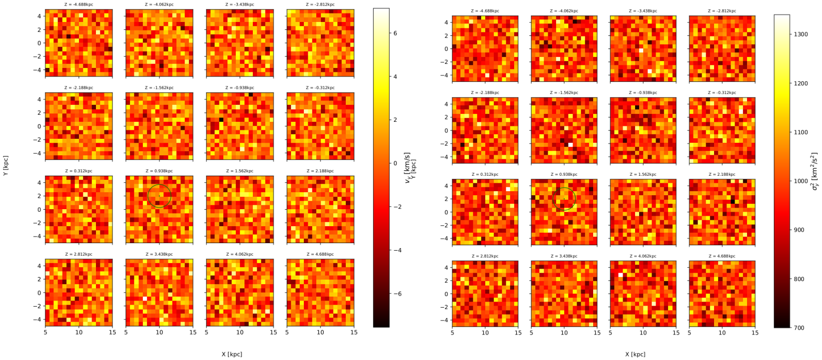


Signatures in stellar field





Signatures in stellar field



Simplified problem: classification

Aim: determine the number of subhalos in given stellar field

Method: Classification via 3D convolutional network

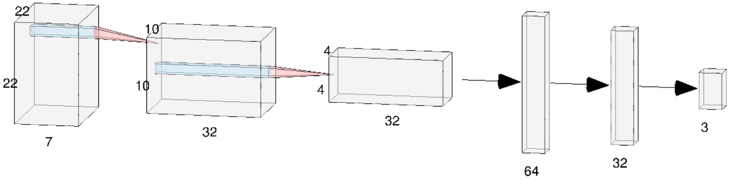


Figure: Simple CNN model build in TensorFlow-Keras.

Preliminary results

Works well for subhalos with $M = 10^7 M_\odot$!

$(n = 10^3\ \mathrm{kpc}^{-3}\ ,\ \sigma_x^2 = \sigma_y^2 = 1000\ \mathrm{km}^2/\mathrm{s}^2$ and $\sigma_z^2 = 100\ \mathrm{km}^2/s^2)$

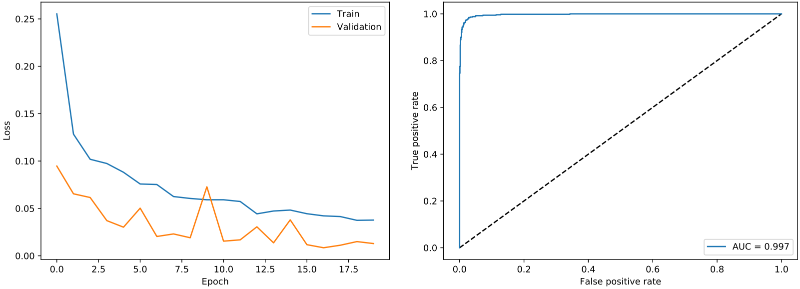


Figure: Cross-entropy loss (left) and ROC curve (right).

Outlook

Not yet ready for realistic data

Many possible **improvements**:

1. Look at subpopulation of stars (cold & numerous)
2. Optimize the observed region (hexagonal convolution?)
3. Consider higher moments of stellar velocity distribution
4. Better data (Gaia DR3/4, VLT, LSST)

More ambitious **goals**:

1. Regression network: extract subhalo masses and velocities
2. Constrain subhalo density profile

Happy to hear your opinion and suggestions!