Seeing the dark matter halo through Gaia's eyes

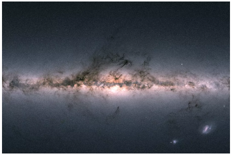
with machine learning

Marat Freytsis

Tel Aviv/IAS

Korea Meeting on Particle Physics
May 3, 2019

Tim Cohen, MF, Mariangela Lisanti, Lina Necib, Bryan Ostdiek
and FIRE collaborations: Shea Garrison-Kimmel, Andrew Wetzel,
Robyn Sanderson, Philip F. Hopkins [arXiv:190i.ntime]



Gaia and dark matter

- Gaia: the largest 5D/6D catalog of local astronomical objects ever
- Can it teach us about the dark matter halo of the Milky Way?
- Why improve our halo models?
  - Astronomers: Learn galactic formation histories
  - Particle physicists: Halo feeds into detection rates
- Older stars act as tracers for (some) dark matter
- The challenge: identifying old stars with Gaia only

Plan

- Gaia and DM
- Halo models and stellar tracers
  - Toy models & merger histories
  - Finding visible tracers of DM
- Machine learning with Gaia through FIRE
  - General methods
  - Validating performance
- Performance in simulation and prospects

Toy models of Milky Way

visible galaxy

Central bulge + disk

us: $\sim 8\,\text{kpc}$ out

$$M_{\text{stellar}} \approx 5 \times 10^{10} M_{\odot}$$

$$z_{\text{disk}} \approx 0.6\,\text{kpc}$$
$$R_{\text{disk}} \approx 15\,\text{kpc}$$
$$R_{\text{bulge}} \approx 4\,\text{kpc}$$

Toy models of the Milky Way
DM halo

rotation curves ($v_c(r) = \sqrt{\frac{GM}{r}}$) $\Longrightarrow$ visible galaxy inside DM halo

$R_{\text{halo}} \sim 100\,\text{kpc}$, $M_{\text{halo}} \sim 10^{12} M_\odot$
flat $v_c(r) \Longrightarrow M(r) \propto r$

$$\rho(r) \propto r^{-2}$$
$$v_c(R_{\text{halo}}) \sim 200\,\text{km}/\,\text{sec}$$

- collisionless
- nonrelativistic
- self-gravitating
- isotropic/isothermal

Toy models of the Milky Way
DM halo

rotation curves ($v_c(r) = \sqrt{\frac{GM}{r}}$) $\Longrightarrow$ visible galaxy inside DM halo

$R_{\text{halo}} \sim 100\,\text{kpc}$, $M_{\text{halo}} \sim 10^{12} M_\odot$
flat $v_c(r) \Longrightarrow M(r) \propto r$

$$\rho(r) \propto r^{-2}$$
$$v_c(R_{\text{halo}}) \sim 200\,\text{km}/\,\text{sec}$$

- collisionless
- nonrelativistic
- self-gravitating
- isotropic/isothermal

Toy models of the Milky Way

DM halo

rotation curves ($v_c(r) = \sqrt{\frac{GM}{r}}$) $\Longrightarrow$ visible galaxy inside DM halo

$R_{\mathrm{halo}} \sim 100\,\mathrm{kpc}$, $M_{\mathrm{halo}} \sim 10^{12} M_\odot$

$$\text{flat } v_c(r) \Longrightarrow \begin{aligned} &M(r) \propto r \\ &\rho(r) \propto r^{-2} \\ &v_c(R_{\mathrm{halo}}) \sim 200\,\mathrm{km}/\sec \end{aligned}$$

- collisionless
- nonrelativistic
- self-gravitating
- isotropic/isothermal

Hierarchical merger model

Where did all this come from?

1. Density fluctuations after big bang lead to protogalactic fragments of $O(10^6\text{–}10^8 M_\odot)$
2. Fragments evolve in isolation creating stars/globular clusters
3. Collisions and tidal disruptions lead to distribution of halo (stars and DM)
4. Gas in the mergers interacts and collapses to disk
5. Young and metal rich stars produced in the disk

The last major merger occurred $\sim 10\,\mathrm{Gyr}$ ago
Minor mergers still happening

Hierarchical merger model

Where did all this come from?

1. Density fluctuations after big bang lead to protogalactic fragments of $O(10^6–10^8 M_\odot)$
2. Fragments evolve in isolation creating stars/globular clusters
3. Collisions and tidal disruptions lead to distribution of halo (stars and DM)
4. Gas in the mergers interacts and collapses to disk
5. Young and metal rich stars produced in the disk

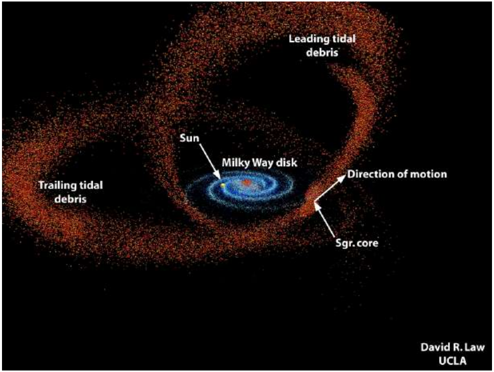
The last major merger occurred $\sim 10\,\mathrm{Gyr}$ ago
Minor mergers still happening

Old stars as tracers

Local halo imprinted with merger history

Stars and DM interact almost only through gravity

To find DM, find stars from early mergers



Tracing DM

How to detect the oldest stars?

Early merger $\longrightarrow$ old star $\longrightarrow$ low metallicity

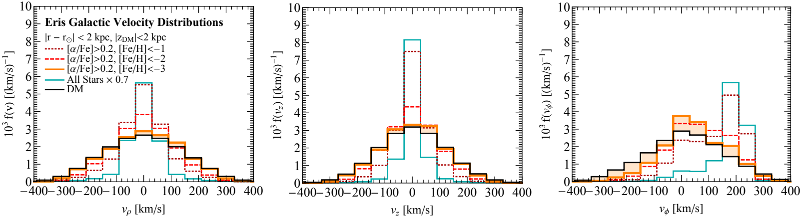
$$[\mathbf{Fe}/\mathbf{H}] = \log_{10}\left(\frac{N_{\mathrm{Fe}}}{N_{\mathrm{H}}}\right) - \log_{10}\left(\frac{N_{\mathrm{Fe}}}{N_{\mathrm{H}}}\right)_{\odot} < C$$

Also helps not to look directly in the disk

$$|z| > z_{\mathrm{cut}}$$

Tracing DM
results in simulation

Does this work?

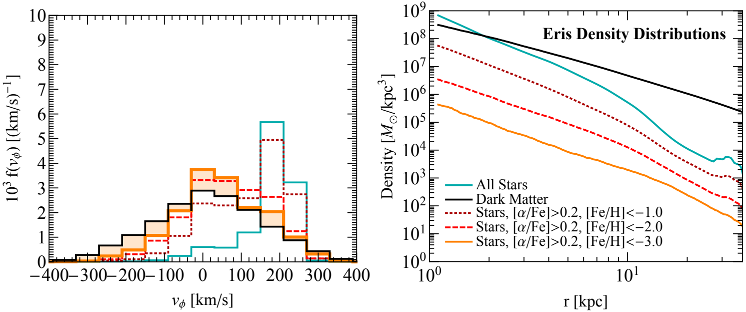


Herzog-Arbeitman, Lisanti, Madau, Necib [arXiv:1704.04499]

Old stars and DM share the same velocity distributions!

Tracing DM

results in simulation

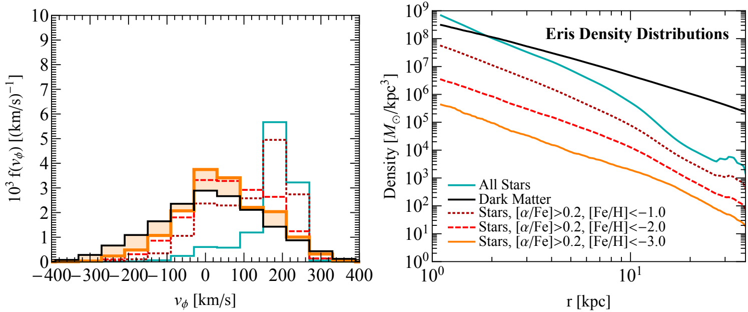


Herzog-Arbeitman, Lisanti, Madau, Necib [arXiv:1704.04499]

Old stars and DM share the same density profile!

Can stellar tracers of virialized DM be isolated in practice?

Tracing DM
results in simulation



Herzog-Arbeitman, Lisanti, Madau, Necib [arXiv:1704.04499]

Old stars and DM share the same density profile!

Can stellar tracers of virialized DM be isolated in practice?

Catalogs of real data

Phase space

- Gaia DR1 (2-D location for 1.1 billion stars)
  - Crossmatched with Hipparcos Tycho-2 catalog (2 million stars)
- Gaia DR2 (5-D PS for 1.3 billion stars)

RAVE-TGAS (255,922 stars)

Spectroscopy + v_r

- RAdial Velocity Experiment
- Sloan Digital Sky Survey

Gaia-SDSS (193,162 stars)

Catalogs of real data

Phase space

- Gaia DR1 (2-D location for 1.1 billion stars)
  - Crossmatched with Hipparcos Tycho-2 catalog (2 million stars)
- Gaia DR2 (5-D PS for 1.3 billion stars)

Spectroscopy + v_r

- RAdial Velocity Experiment
- Sloan Digital Sky Survey

RAVE-TGAS (255,922 stars)

Gaia-SDSS (193,162 stars)

Catalogs of real data

Phase space

- Gaia DR1 (2-D location for 1.1 billion stars)
  - Crossmatched with Hipparcos Tycho-2 catalog (2 million stars)
- Gaia DR2 (5-D PS for 1.3 billion stars)

RAVE-TGAS (255,922 stars)

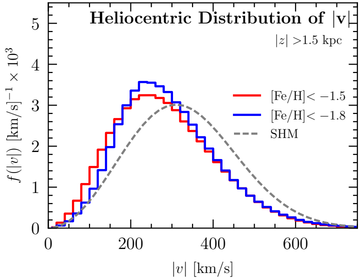
Spectroscopy + v_r

- RAdial Velocity Experiment
- Sloan Digital Sky Survey

Gaia-SDSS (193,162 stars)

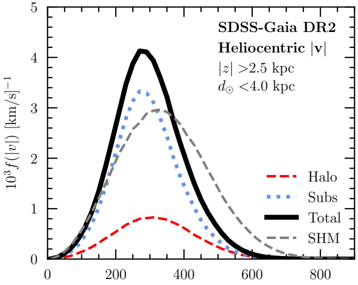
…and real-world results

RAVE-TGAS



[arXiv:1708.03635]

Gaia-SDSS



[arXiv:1807.02519]

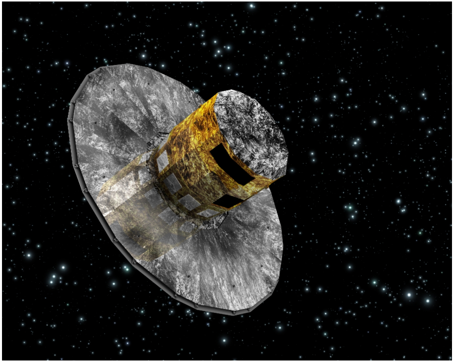
virialized DM velocities smaller than standard halo model
$\Longrightarrow$ implications for DM direct detection

But accuracy limited by cross-correlating data

Plan

- Gaia and DM
- Halo models and stellar tracers
  - Toy models & merger histories
  - Finding visible tracers of DM
- Machine learning with Gaia through FIRE
  - General methods
  - Validating performance
- Performance in simulation and prospects

Letting Gaia see on its own



DR2: 5-D kinematics and 2-band spectroscopy on 1.3 billion stars

Not enough information to extract metallicity conventionally
Idea: Use neural network classifier as old star distribution fitter

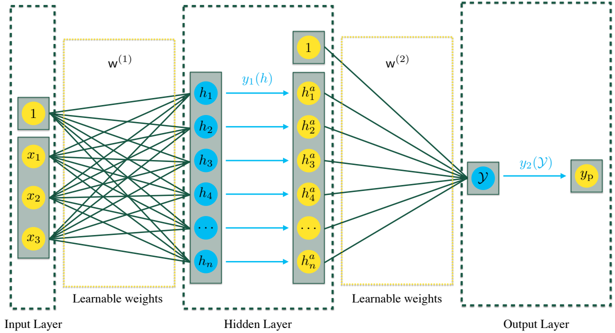
Gaia data format details

Stellar information provided

- Galactic longitude and latitude (ℓ, b)
- Proper motion in right ascension and declination $(\mu_{\alpha,\delta})$
- Parallax
- Blue- and red-band magnitude $(G_{\mathrm{BP,RP}})$

Provides 5D phase-space information (radial v missing)
Complementary information to parallax in G
if neural network can learn distance–luminosity function
Residual information about metallicity also in G?

Feed-forward NN classification



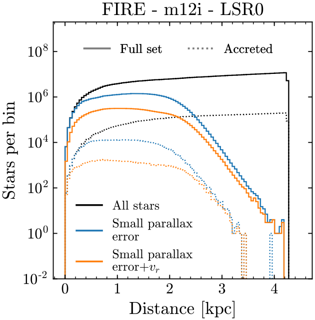
$$\ell_{\mathrm{BCE}}(\{y_t\}, \{y_p\}) = -\sum_i \left(y_{t,i} \log y_{p,i} + (1 - y_{t,i}) \log(1 - y_{p,i})\right)$$

Requires event-by-event labels for (simulated) training sample

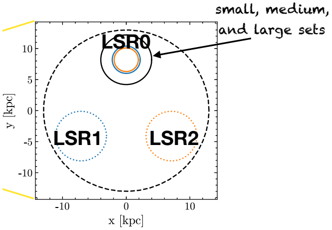
- Use FIRE simulations with labels from known history

Network and training procedure

- Train 5-layer network
  - 7 inputs à la Gaia
  - 3 hidden layers of 100 nodes each
  - star classified as accreted or not
- Label from FIRE merger history
  - Remove metallicity middleman
- 600 million stars per viewpoint
- Include measurement uncertainty by resampling each star within its errors 20 times



Crosschecks and transfer learning

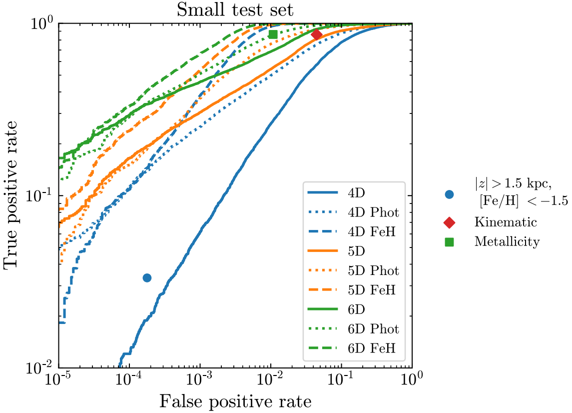


- Maybe just learn particular local distribution/merger history?
  - Compare different observations
  - Compare different simulations
- Systematic errors in FIRE mocks?
- Compensate via transfer learning
  - Lower NN layers learn simple cuts
  - High-level observables in top layer
  - Train full network on a dataset
  - Reset *top layer only* and retrain *only that layer* on new data
  - Requires much less data in 2nd set
  - Reduce sensitivity to complex features in original training set

Plan

- Gaia and DM
- Halo models and stellar tracers
  - Toy models & merger histories
  - Finding visible tracers of DM
- Machine learning with Gaia through FIRE
  - General methods
  - Validating performance
- Performance in simulation and prospects
  - All results preliminary!

Classifying close stars

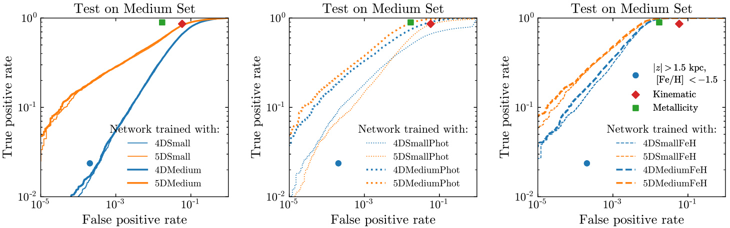


Close stars have multiple parallax measurements

$\longrightarrow$ radial velocity recovered, full 6-D PS information available

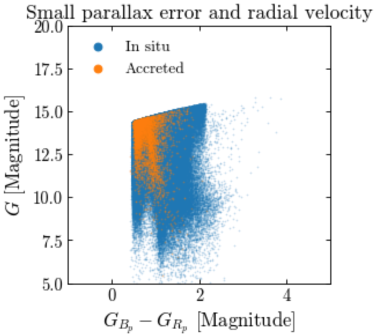
Photometric data help when only reduced PS information exists

Moving farther out



At best only 5-D information at larger differences
5-D information or photometric data critical to best performance

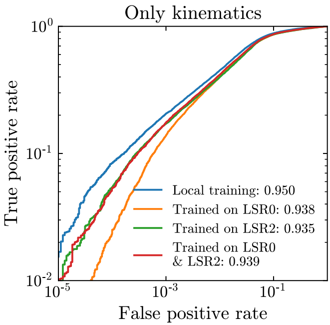
A closer look at photometric data

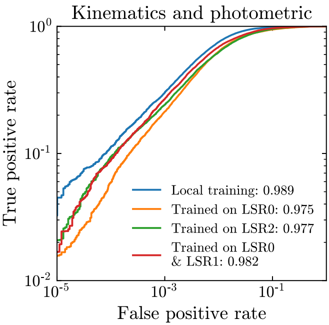


At smaller distances, training data doesn't cover full HR diagram
Luminocity-distance relations not fully learned

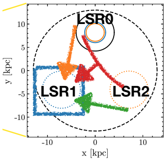
must be careful training set goes out as far as real data with photometry

Comparing viewpoints
Testing on LSR1



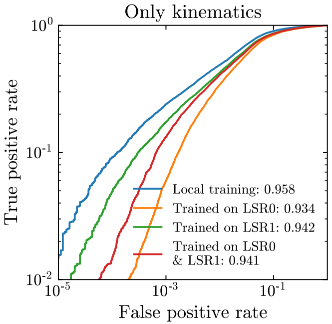


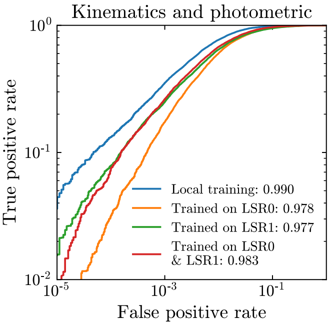
training on multiple viewpoints
$\Longrightarrow$ improved generalization



Comparing viewpoints

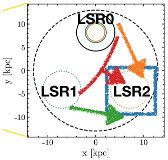
Testing on LSR2





details depend on local kinematics

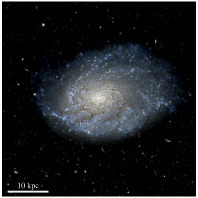
seemingly more stable generalization with $G_{\mathrm{BP,RP}}$

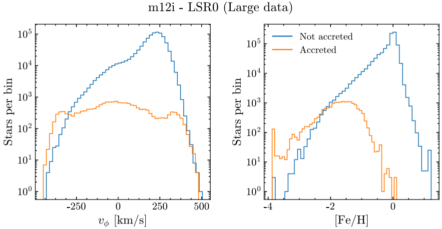


Trying a new galaxy

old galaxy

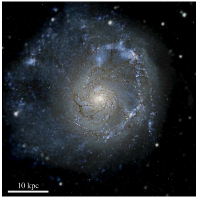
m12i

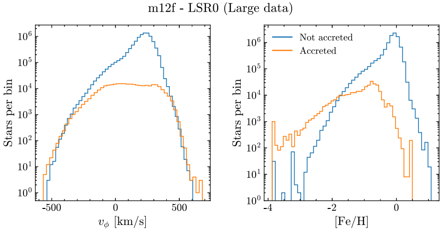




new galaxy

m12f





Different merger history indicated by v_φ distribution

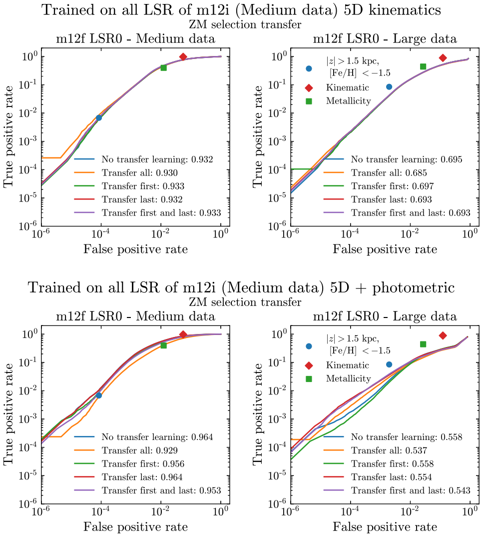
Transfer learning results

Medium set

More realistic scenario: labels after transfer from cuts, not truth

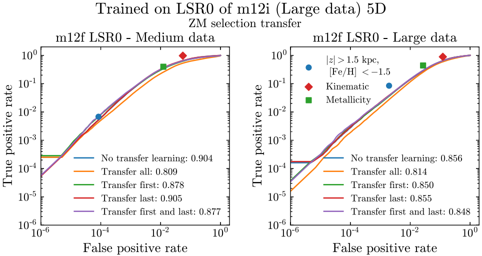
Use $|z|$ and [Fe/H] to set labels in m12f

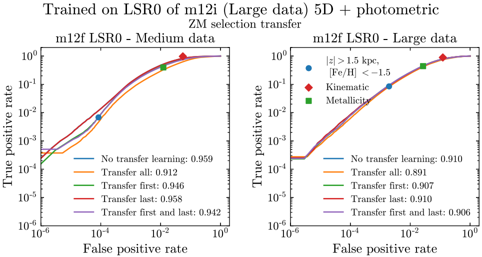
Most effective to recalibrate $G_{\mathrm{BP,RP}}$ and large-distance kinematics



Transfer learning results
Large set

With big training set, transfer less necessary



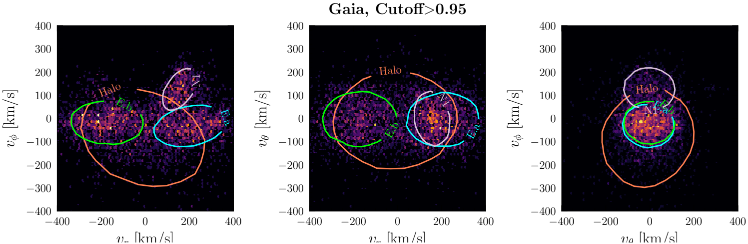


Preliminary look at Gaia DR2

Preliminary!

Have created an (expected) 95% pure accreted Gaia DR2 dataset

- Contains 21304 stars with full 6D information



Gaia-Enceladus clearly visible
New structure (Nyx) visible without rotational symmetry
Will introduce asymmetry to local DM velocity if confirmed

Next steps

- Final validation and closure tests in FIRE Gaia mocks
- Release a public catalog of virialized old stars in Gaia dataset
- Extraction of local DM halo v distributions
  - Effect on DM direct detection rates
  - Characterize uncertainties in the method
- Can be say anything about unvirialized/unresolved DM?

Conclusions

- Hierarchical mergers imply old stars are efficient DM tracers
  - metallicity and kinematics serve as efficient selection criteria
  - Gaia has no access to metallicity; cut-based analyses insufficient
- Modern machine learning techniques allow the full resolving power of the Gaia dataset to be brought on the problem
  - Kinematic and spectral information can be as powerful
  - Training must be performed carefully to avoid sample bias
  - Transfer learning techniques help control systematics
- Validation in simulation nearly complete
  - Real data calalogs and physics analyses released soon
  - New local structures in the glaxay already located
  - Correlations of structures with DM affect DD interpretations
- ML gives a path to unlocking the full potential of the Gaia

Thank you!