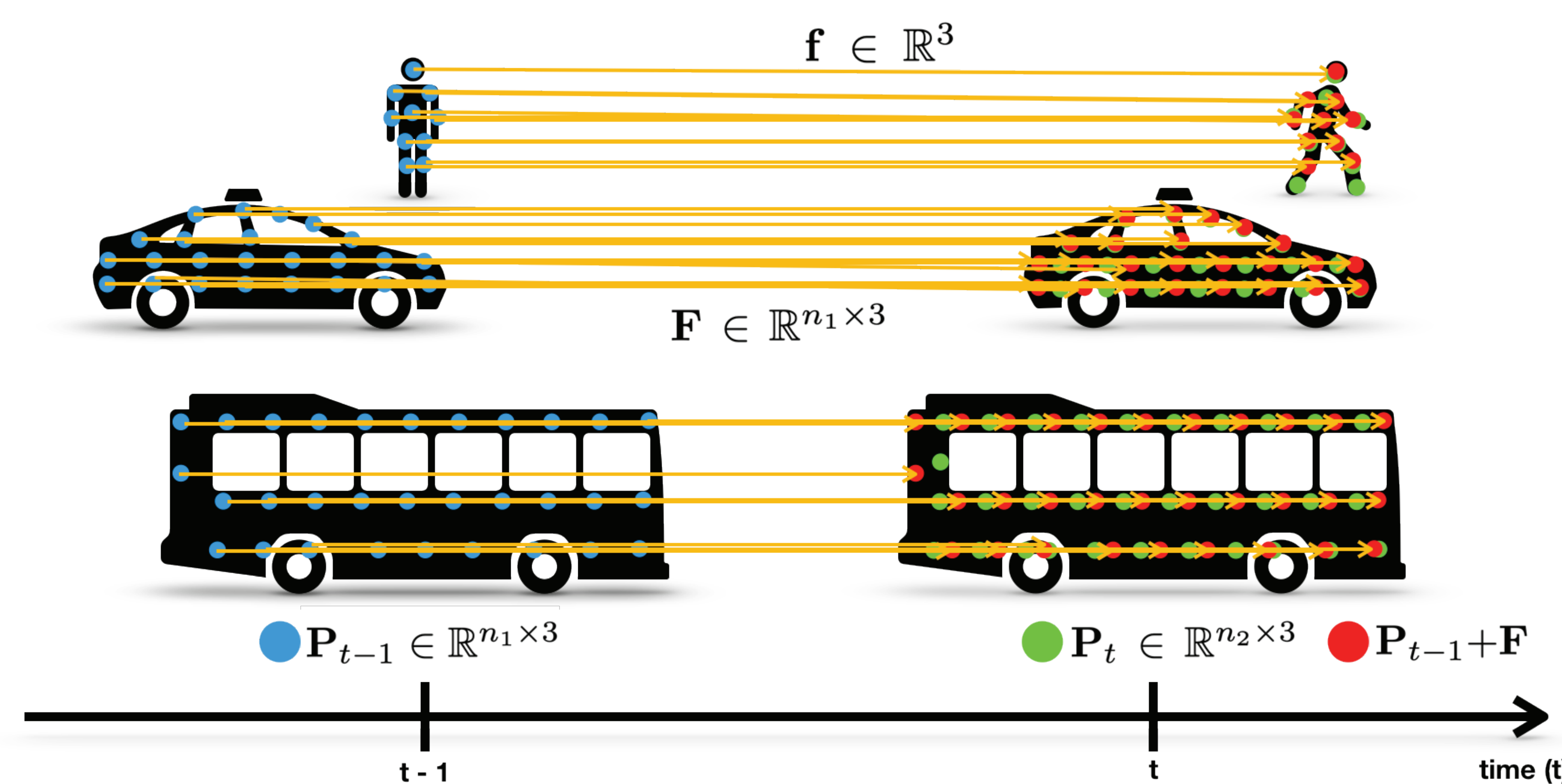# Scene Flow from Point Clouds with or without Learning

Jhony Kaesemodel Pontes[1]   James Hays[1,2]   Simon Lucey[3]

[1]Argo AI   [2]Georgia Institute of Technology   [3]The University of Adelaide

What is scene flow?

$\mathbf{f} \in \mathbb{R}^3$

$\mathbf{F} \in \mathbb{R}^{n_1 \times 3}$

$\mathbf{P}_{t-1} \in \mathbb{R}^{n_1 \times 3}$ $\mathbf{P}_t \in \mathbb{R}^{n_2 \times 3}$ $\mathbf{P}_{t-1} + \mathbf{F}$

t - 1   t   time (t)
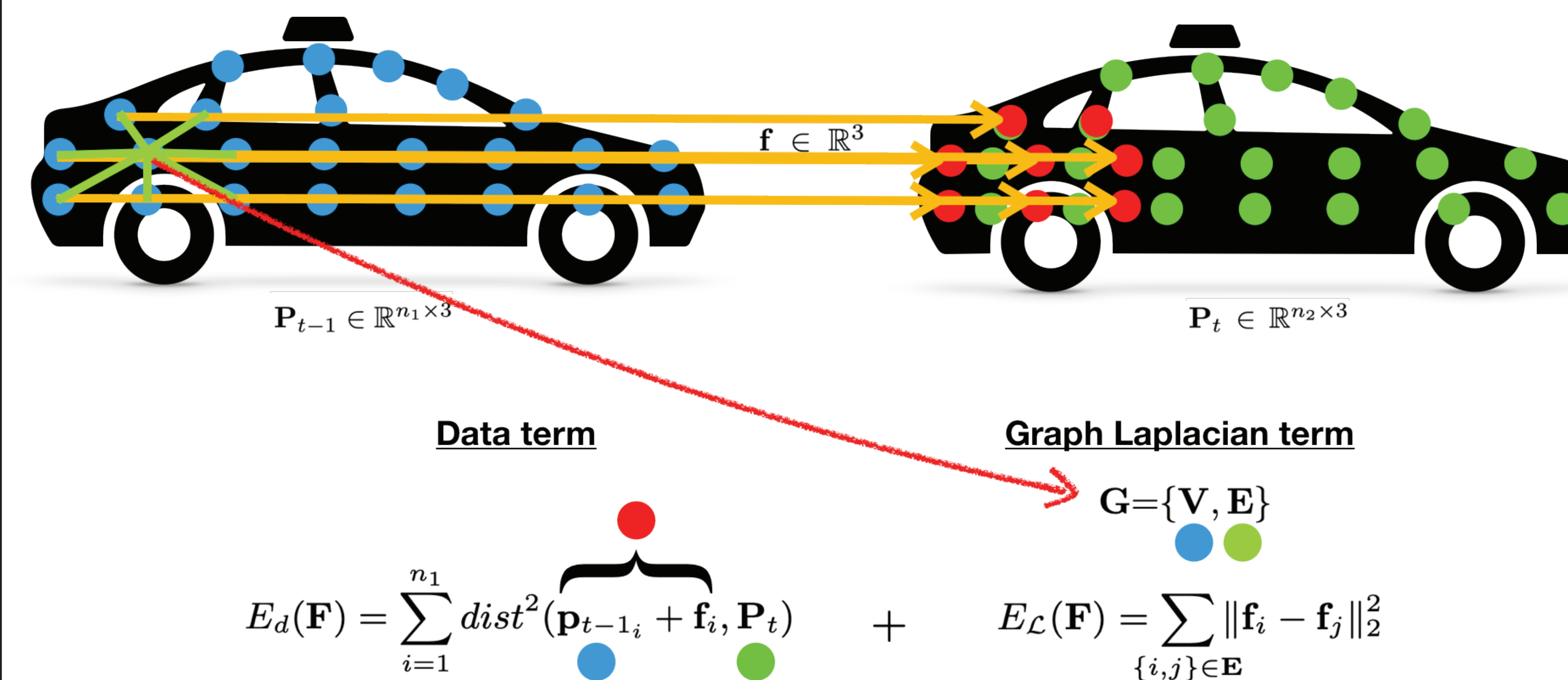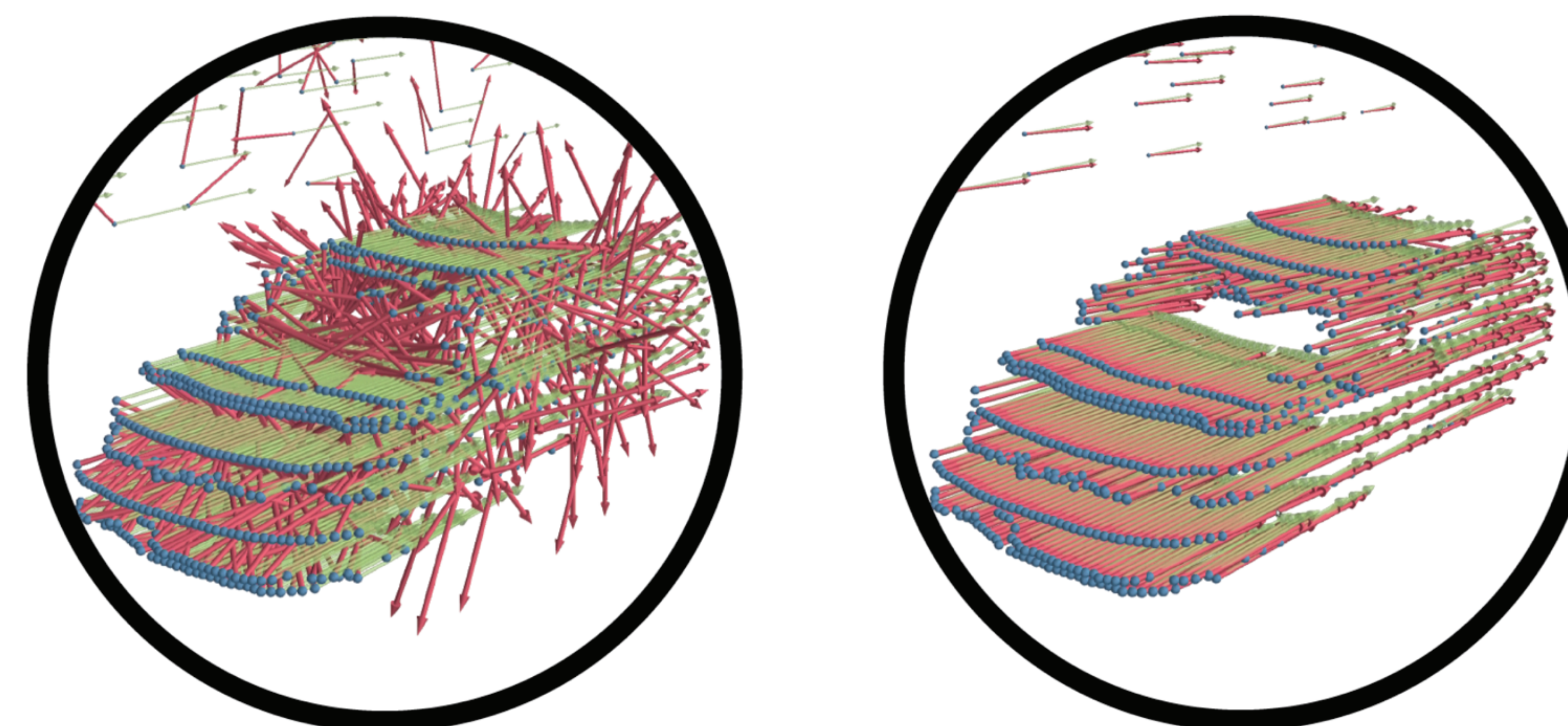
Our motivations

- Learning scene flow demands annotated data
- Real-world scene flow annotations are scarce
- Learning-based methods are domain specific
- Inspiration from recent self-supervised works
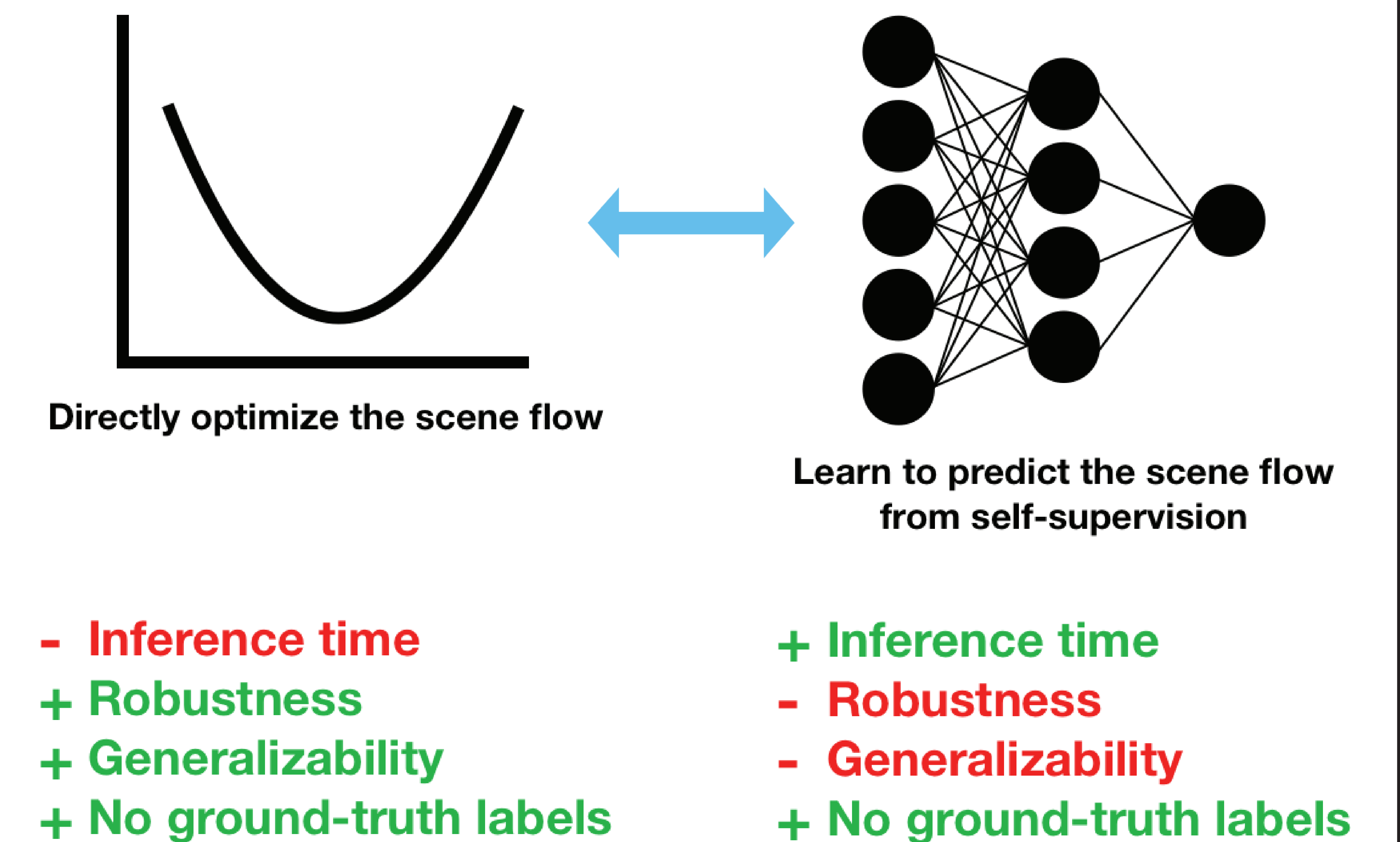- Inspiration from classical non-learning works

How do we estimate it?

$\mathbf{f} \in \mathbb{R}^3$

$\mathbf{P}_{t-1} \in \mathbb{R}^{n_1 \times 3}$ $\mathbf{P}_t \in \mathbb{R}^{n_2 \times 3}$

Data term

$$E_d(\mathbf{F}) = \sum_{i=1}^{n_1} dist^2(\mathbf{p}_{t-1_i} + \mathbf{f}_i, \mathbf{P}_t)$$

Graph Laplacian term

$\mathbf{G} = \{\mathbf{V}, \mathbf{E}\}$

$$E_{\mathcal{L}}(\mathbf{F}) = \sum_{\{i,j\} \in \mathbf{E}} \|\mathbf{f}_i - \mathbf{f}_j\|_2^2$$

Why the graph Laplacian term?

(a) Without the graph Laplacian   (b) With the graph Laplacian

With or without learning?

Directly optimize the scene flow

Learn to predict the scene flow from self-supervision

− Inference time	+ Inference time
+ Robustness	− Robustness
+ Generalizability	− Generalizability
+ No ground-truth labels	+ No ground-truth labels

Visual result

$\mathbf{P}_{t-1} \in \mathbb{R}^{n_1 \times 3}$
$\mathbf{P}_t \in \mathbb{R}^{n_2 \times 3}$
$\mathbf{P}_{t-1} + \mathbf{F}$

0.0 0.3 0.6 0.8 1.1 1.4 1.7 2.0 2.3 2.5