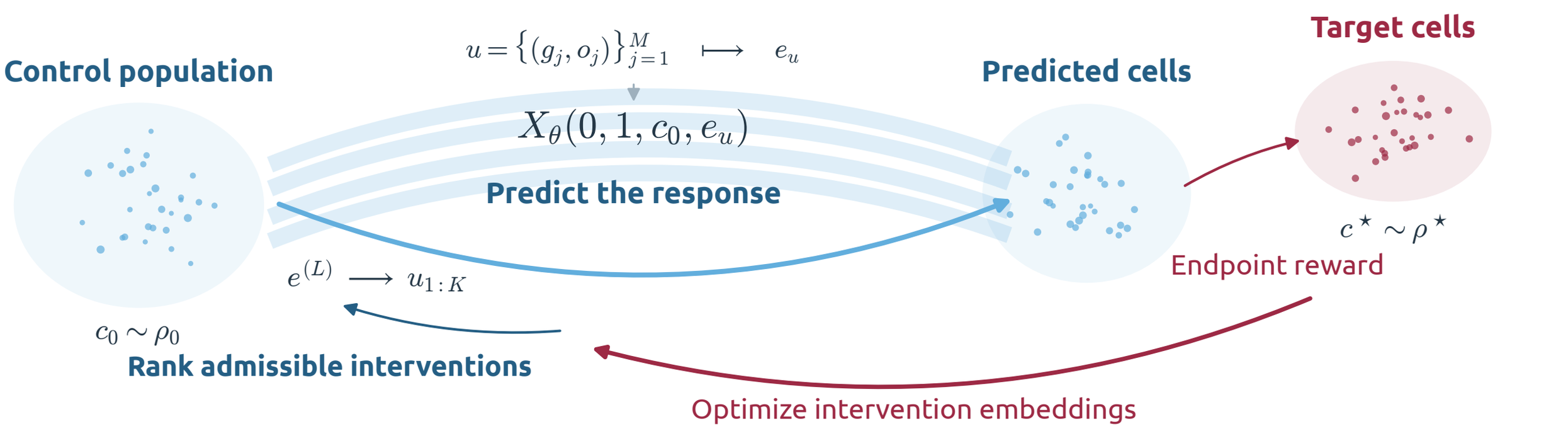


PIVOT

$$\hat{c}_1 = X_\theta(0, 1, c_0, e) \qquad g_e = J_e X_\theta^\top \nabla_{\hat{c}_1} r \qquad e^+ = e + \gamma \frac{g_e}{\|g_e\|_2 + \epsilon}$$

Endpoint prediction, reward gradients, and admissible interventions.