which directly approximates the marginal distributions of continuous robotic states via a collection of M sample points. Messages are represented via a weighted mixture of M Gaussian distributions:

$$m_{st}(x_t) = \sum_{i=1}^{M} w_{st}^{(i)} \mathbb{N}(x_t \mid x_{st}^{(i)}, \Sigma_{st}) \tag{9}$$

Mixture components are centered on samples $x_{st}^{(i)}$ from the underlying, continuous message function, with weights $w_{st}^{(i)}$ set via importance sampling principles as detailed below. By choosing a number of samples M, we can tradeoff accuracy versus computational cost.

A variety of methods are available for choosing the covariance Σ_{st} used to smooth message samples [29]. In many cases, we use the computationally efficient "rule of thumb" estimate $\Sigma_{st} = \mathrm{ROT}(\{x_{st}^{(i)}, w_{st}^{(i)}\})$, which is proportional to the weighted covariance of the observed samples:

$$\Sigma_{st} = M^{\frac{-2}{\delta+4}} \sum_{i=1}^{M} w_{st}^{(i)} (x_{st}^{(i)} - \bar{x}_{st})(x_{st}^{(i)} - \bar{x}_{st})^T \tag{10}$$

Here, $\bar{x}_{st} = \sum_i w_{st}^{(i)} x_{st}^{(i)}$ is the weighted sample mean, and $\delta = \dim(x_t)$. However, for ring-shaped messages which are far from unimodal, the rule of thumb estimator performs poorly, significantly oversmoothing the estimated density. In such cases, we set $\Sigma_{st} = \sigma_\nu^2 \xi_M I$, where ξ_M is a constant calibrated offline to the number of samples M.

D. Decomposing Gaussian Mixtures

In the tracking algorithm developed in Sec. VI, it is necessary to decompose a Gaussian mixture $p(\chi) = p(x, \dot{x})$ into marginal and conditional distributions $p(x)p(\dot{x} \mid x)$. Because the conditional distribution of a subset of jointly Gaussian variables is Gaussian, conditional distributions of Gaussian mixtures are mixtures of lower-dimensional Gaussians. We now derive formulas for the marginal and conditional portions of each mixture component. As the same formulas apply to all components, we drop the $\cdot^{(i)}$ notation.

Consider a Gaussian mixture component with weight w_χ, and mean vector and covariance matrix

$$\mu_\chi = \begin{bmatrix} \mu_x \\ \mu_{\dot{x}} \end{bmatrix}, \qquad \Sigma_\chi = \begin{bmatrix} \Sigma_{x,x} & \Sigma_{x,\dot{x}} \\ \Sigma_{\dot{x},x} & \Sigma_{\dot{x},\dot{x}} \end{bmatrix}. \tag{11}$$

The marginal $p(x)$ has weight w_χ, mean μ_x, and covariance matrix Σ_x. The conditional density $p(\dot{x} \mid x = a)$ then has the following mean, covariance, and weight:

$$\mu_{\dot{x}|x} = \mu_{\dot{x}} + \Sigma_{\dot{x},x}\Sigma_{x,x}^{-1}(a - \mu_x) \tag{12}$$

$$\Sigma_{\dot{x}|x} = \Sigma_{\dot{x},\dot{x}} - \Sigma_{\dot{x},x}\Sigma_{x,x}^{-1}\Sigma_{x,\dot{x}} \tag{13}$$

$$w_{\dot{x}|x} \propto w_\chi \left(\frac{|\Sigma_{\dot{x}|x}|}{|\Sigma_\chi|}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\sigma_0^T \Sigma_\chi^{-1} \sigma_0\right\} \tag{14}$$

The transformations from a single joint multivariate Gaussian to a single conditional multivariate Gaussian, transforming the mean and standard deviation, are determined by standard formulas[30]. The weight $w_{\dot{x}|x}$ depends on a centered observation vector:

$$\sigma_0 = \begin{bmatrix} a - \mu_x \\ \mu_{\dot{x}|x} - \mu_{\dot{x}} \end{bmatrix} \tag{15}$$

We use this decomposition to factorize dynamics messages sent between the same robot at subsequent timesteps. By first integrating distance readings via the smaller position-only marginals, and then drawing appropriately reweighted velocity samples, we can integrate multi-hop inter-distance information over time.

V. Phase I: Localization Algorithm

The goal of the inter-distance localization sub-problem is to infer the location of each node using only information from a single timestep (the first). Our formulation closely follows the algorithm proposed by Ihler et al. [6]. Since we consider a single timestep, we drop the τ notation and velocity state variables, and denote the position of robot s by x_s. The NBP algorithm alternates between sending messages to neighboring nodes, and computing marginals at nodes which received messages. We refer to each pair of steps, in which all messages are updated once, as an iteration.

To compute new outgoing messages, we use the marginal estimate $\hat{p}_s^{n-1}(x_s)$ from the previous iteration, which is represented by a set of samples. The inter-distance sensor model of Eq. (2) implies that message samples are most easily generated in polar coordinates, with random orientation and radius perturbed from the noisy inter-distance reading:

$$\theta_{st}^{(i)} \sim \mathbb{U}(0, 2\pi) \qquad \nu_{st}^{(i)} \sim \mathbb{N}\left(0, \sigma_\nu^2\right) \tag{16}$$

$$x_{st}^{(i)} = x_s^{(i)} + (d_{st} + \nu_{st}^{(i)})[\sin(\theta_{st}^{(i)});\ \cos(\theta_{st}^{(i)})] \tag{17}$$

When the BP algorithm sends a message from s to t, the incoming message product (as in Eq. (7)) avoids double-counting by excluding the message sent from t to s. To account for this in our NBP implementation, we use an importance sampling framework [4], [6]. Each sample $x_{st}^{(i)}$ is reweighted by $1/m_{ts}^{n-1}(x_s^{(i)})$, where $m_{ts}^{n-1}(x_s)$ is the preceding iteration's message (a Gaussian mixture). Also accounting for the effects of distance on the probability of receiving a measurement, the overall sample weight is

$$w_{st}^{(i)} = P_o(x_{st}^{(i)})/m_{ts}^{n-1}(x_s^{(i)}). \tag{18}$$

More generally, given a proposal density $q(x)$ from which we can draw samples $x^{(i)}$, and a target density $p(x)$ we would like to approximate, importance sampling methods [31] assign a weight $w^{(i)} \propto p(x^{(i)})/q(x^{(i)})$ to each sample $x^{(i)}$.

Importance sampling methods are also used to approximate the marginal update of Eq. (8). Recall that incoming messages $m_{ts}(x_s)$ are mixtures of M Gaussians. With $d = |\Gamma(s)|$ neighbors, the exact message product of Eq. (8) will produce an intractable mixture with M^d components. To construct a proposal distribution, we instead take an equal number of samples $x_s^{(i)} \sim m_{ts}(x_s)$ from each incoming message. Combining these d groups of samples into a single