CEA Comité Español de Automática
XLVII JORNADAS DE AUTOMÁTICA CÓRDOBA 2026
UNIVERSIDADE DA CORUÑA

Jornadas de Automática

Robust Task Generalization for Dual-Arm Learning from Demonstration

Prados, Adrian [a,*], Lishan, Lucia [a], Mendez, Alberto [a], Garrido, Santiago [a], Barber, Ramon [a]

[a] *RoboticsLab, Departamento de Ingeniería de Sistemas y Automática, Universidad Carlos III de Madrid, Av. Universidad, nº30, 28911, Leganés, España.*

To cite this article: Prados, A, Lishan, L, Mendez, A, Garrido, S, Barber, R. 2026. Robust Task Generalization for Dual-Arm Learning from Demonstration. Jornadas de Automática, 47.
https://doi.org/10.17979/ja-cea.2026.47.13706

Resumen

La manipulación bibrazo o la coordinación física humano-robot, requieren que los robots se adapten rápidamente a entornos y restricciones cambiantes. Los enfoques tradicionales de Aprendizaje por Demostración tienen dificultades para generalizar frente a escenarios fuera de distribución, requiriendo costosos reentrenamientos. Proponemos un algoritmo de aprendizaje de Primitivas de Movimiento basado en Procesos Gaussianos, combinado con una adaptación zero-shot en tiempo real mediante Pathwise Conditioning. El método encapsula la incertidumbre predictiva del movimiento demostrado mediante GPs heterocedásticos y utiliza una actualización mediante la regla de Matheron para ajustar la trayectoria a nuevos via-points de manera instantánea, sin necesidad de reentrenar el modelo subyacente. Esta formulación se extiende a la coordinación bimanual calculando dinámicamente restricciones relativas en 6D para mantener una cadena cinemática cerrada. Los resultados experimentales, tanto en comparativas 2D contra modelos parametrizados por tareas como en tareas con el robot ADAM, demuestran una adaptación robusta con un error cercano a cero y en tiempo real, pudiendo aplicarse para entornos altamente cambiantes.

Palabras clave: Aprendizaje por Demostración, Manipulación bibrazo, Aprendizaje Adaptativo.

Abstract

Dual-arm manipulation or physical human-robot coordination requires robots to adapt rapidly to changing environments and constraints. Traditional Learning from Demonstration approaches struggle to generalize when faced with out-of-distribution scenarios, requiring costly retraining. We propose a Movement Primitive learning algorithm based on Gaussian Processes, combined with real-time zero-shot adaptation through Pathwise Conditioning. The method encapsulates the predictive uncertainty of the demonstrated movement using heteroscedastic GPs and utilizes an update via Matheron's rule to instantaneously adjust the trajectory to new via-points, without the need to retrain the underlying model. This formulation is extended to dual-arm coordination by dynamically calculating 6D relative constraints to maintain a closed kinematic chain. Experimental results, both in 2D comparisons against task-parameterized models and in tasks with the ADAM robot, demonstrate robust adaptation with near-zero error in real time, making it applicable for highly changing environments.

Keywords: Learning from Demonstration, Dual-arm manipulation, Adaptative Learning.

1. Introduction

Robotics has been increasingly introduced into everyday environments (Barber et al., 2022), which are dynamic and unstructured. To operate effectively in these human-centered environments, robots must possess the capability to adapt rapidly to novel and changing situations. This becomes critical in tasks requiring complex physical interaction, such as coordinated dual-arm manipulation (transporting boxes) or physical human-robot interaction (pHRI), where the robot and the human share control over an object. Endowing robots with these skills through explicit programming is extremely complex, which has motivated the use of techniques such as Learning from Demonstration (LfD) (Prados et al., 2023, 2026a). Although LfD provides a powerful framework for transferring human expertise to robots, many approaches exhibit limitations when generalizing beyond the set of demonstrations,

*Correspondence author: aprados@ing.uc3m.es

constraining their ability to adapt to unforeseen scenarios. This vulnerability to out-of-distribution (Spencer et al., 2021) situations is particularly problematic in collaborative tasks. In dual-arm manipulation or human-robot cooperation, the environment is inherently stochastic; the behavior of the human partner or the position of the targets can change mid-execution. Therefore, the robot should not merely memorize trajectories for specific configurations, but rather continuously adapt its knowledge to new spatial constraints while preserving the structure of the original skill.

Among the probabilistic LfD approaches developed to address task generalization, there are methods based on data variability, such as Gaussian Mixture Models (GMM) (Feng et al., 2025) or Probabilistic Movement Primitives (ProMP) (Ruan et al., 2024), and methods based on prediction uncertainty, such as Gaussian Process Regression (GPR) (Arduengo et al., 2021). However, most of these techniques do not consider both metrics (variability and data availability) jointly, and they require computationally expensive retraining processes or multi-context demonstrations to generalize properly. This precludes reactive, real-time adaptation, which is essential when an instantaneous trajectory adjustment is required due to the via-point modifications exerted by a human or changes in the POSE of the task parameters.

In this work we propose the application of a Movement Primitive (MP) learning algorithm based on Gaussian Processes (GPs), combined with a zero-shot adaptation to new via-points through Pathwise Conditioning. Unlike traditional approaches that require recalculating the entire regression or relying on synthetic data, this method allows for real-time trajectory correction by capturing both uncertainty and variability. By leveraging this probabilistic framework, the algorithm not only learns the movement policy from a single demonstration but is also capable of adapting it rapidly while preserving prior knowledge. In this paper, we apply and extend this mathematical formulation to solve dual-arm coordination and human-robot collaboration tasks, enabling the robotic system to dynamically modify its start, via-, and end points in real time—in both position and orientation—in response to changes imposed by the environment or a human operator.

2. Related Work

Various LfD approaches exist to address the out-of-distribution problem (Mehta et al., 2025). Probabilistic LfD methods have emerged as some of the most prominent within this domain. Several studies have focused on generalization through multitask learning (Mendez et al., 2024; Prados et al., 2024b), learning a shared policy across related tasks by leveraging common information. While these methods enable the adaptation of knowledge to novel situations, they require substantial amounts of data and extensive training procedures.

In recent years, generalization efforts have shifted towards environment observation. Techniques such as *task parametrization* (Calinon, 2016) structure tasks as a series of via-points or reference frames (task frames), acting as contextual elements. This ensures that trajectories are not exclusively encoded in the robot's global coordinates, providing each task with a local representation. Task parameterization has been widely adopted in methods such as Task-Parameterized Gaussian Mixture Models (TP-GMM) (Prados et al., 2024a; Carrasco et al., 2024), Task-Parameterized Dynamic Movement Primitives (TP-DMP) (Huang et al., 2025), and Task-Parameterized Gaussian Processes (TP-GP) (Arduengo et al., 2021). Although generally effective, these methods exhibit limitations when dealing with sparse demonstrations, being highly sensitive to the selection and placement of task parameters, struggling when faced with redundant or rapidly changing elements. Alternative techniques focus directly on learning motion policies. Their objective is to explicitly adapt the policy parameters without the need for external reference elements. Prominent among these are approaches based on ProMPs (Paraschos et al., 2013), which adapt the policy descriptors to conform to new target points.

These algorithms have also been applied to task generalization in coordinated dual-arm scenarios. (Lee et al., 2025) presents a GMM-based approach where box transportation tasks are adapted through the reparametrization of the Gaussians comprising the demonstrations. Silvério et al. (2015) applies task parametrization to an everyday activity, such as sweeping, which requires the coordination and cooperation of both arms. Prados et al. (2025b) introduced an extension of TP-GMM for the coordination of dual-arm manipulation tasks, ensuring collision avoidance between two arms. In addition to this, efforts have been made to apply generalization to human-robot coordination tasks. Notable among these are GMM/GMR-based techniques (Sosa-Ceron et al., 2022), which employ a two-phase learning approach that switches the leader-follower strategy in each phase, as well as ProMP-based approaches (Qian et al., 2022), which probabilistically model the human-robot correlation, thereby enabling the learning of distinct transportation patterns.

3. Learning via Gaussian Processes

3.1. Gaussian Process for Tasks Learning

Gaussian Processes provide a non-parametric probabilistic framework to model latent functions $f(t)$ from noisy demonstrations (Prados et al., 2026b). GPs are defined by its mean function $m(t)$ and covariance function $k(t, t')$, such that $f(t) \sim \mathcal{GP}(m(t), k(t, t'))$.

Given a dataset of demonstrations $\mathcal{D} = \{(t_k, q_k)\}_{k=1}^{N}$, where t_k is the time and $q_k = h(t_k) + \epsilon$ is the observed state with Gaussian noise $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$, the joint distribution of the observed targets and latent values at test inputs t^* is Gaussian. The predictive posterior distribution for $h(t^*) \in \mathbb{R}^M$ is given by $p(h(t^*) \mid t^*, \mathcal{D}) = \mathcal{N}(\mu(t^*), \Sigma(t^*))$, where:

$$\begin{cases} \mu(t^*) = \mathbf{m}(t^*) + K(t^*, t)(K(t, t) + \sigma_n^2 I)^{-1}(q - \mathbf{m}(t)), \\ \Sigma(t^*) = K(t^*, t^*) - K(t^*, t)(K(t, t) + \sigma_n^2 I)^{-1} K(t, t^*). \end{cases} \tag{1}$$

In this work, the kernel $k(\cdot, \cdot)$ hyperparameters are optimized by maximizing the log-marginal likelihood, automated via Approximate Bayesian Computation (ABC) (Arduengo et al., 2021) to ensure robust trajectory encapsulation.

3.2. Pathwise Conditioning for Zero-Shot Adaptation

While standard GPs characterize conditional distributions analytically, calculating the full posterior at each control step is computationally prohibitive for real-time multi-agent coordination. Pathwise conditioning (Wilson et al., 2021) offers

an efficient alternative by sampling before conditioning, decoupling the prior from the local data updates. The mathematical foundation of this approach relies on Matheron's Update Rule. For jointly Gaussian variables, the conditional random variable $\mathbf{a}$ given $\mathbf{b} = \boldsymbol{\beta}$ can be expressed as:

$$(\mathbf{a} \mid \mathbf{b} = \boldsymbol{\beta}) \overset{d}{=} \mathbf{a} + \Sigma_{\mathbf{a},\mathbf{b}}\Sigma_{\mathbf{b},\mathbf{b}}^{-1}(\boldsymbol{\beta} - \mathbf{b}), \tag{2}$$

where $\mathbf{a}$ and $\mathbf{b}$ are unconditional draws from the prior, and the second term is a deterministic update enforcing the condition.

Applied to GPs modelling robotic trajectories $f \sim \mathcal{GP}(0, k)$, let $\mathbf{X}_n$ be a set of via-points (e.g., dynamic constraints imposed by a human or a second robotic arm) taking values $\mathbf{y}$. The pathwise conditioned trajectory is computed as:

$$f(\cdot) \mid f(\mathbf{X}_n) = \mathbf{y} \overset{d}{=} f(\cdot) + k(\cdot, \mathbf{X}_n)K_n^{-1}(\mathbf{y} - f(\mathbf{X}_n)). \tag{3}$$

This formulation is crucial for our collaborative framework: the term $f(\cdot)$ represents the nominal learned movement primitive, while the deterministic term $k(\cdot, \mathbf{X}_n)K_n^{-1}(\mathbf{y}-f(\mathbf{X}_n))$ acts as a real-time trajectory corrector. It enforces the new kinematic constraints ($\mathbf{y}$) instantaneously at 30Hz without retraining the underlying GP model, preserving the shape and intent of the original demonstration.

4. Collaborative LfD via Pathwise Conditioning

We present a unified algorithm for real time task learning and adaptation. First, a LfD phase encapsulates movement primitives and their predictive uncertainty using GPs; second, a zero-shot adaptation phase enforces new constraints continuously through Pathwise Conditioning. This yields highly reactive policies that exploit prior knowledge without the computational burden of retraining, making it useful for closed-chain kinematic constraints required in collaborative manipulation.

4.1. Movement Primitive Encapsulation via GPs

To encapsulate MPs in a data-driven, non-parametric manner, we employ our own heteroscedastic GP formulation (see Fig. 1). This encodes predictive uncertainty at each time step, facilitating confidence-dependent control decisions while allowing the strict enforcement of initial via-point constraints.

I) Filtering and Temporal Alignment. From the acquired kinesthetic data y, a dynamic envelope (defined by bounds lb, up) is estimated to exclude outliers, retaining the region of interest $y_{\text{in}} = \{y_i \mid lb_i \leq y_i \leq up_i, \forall i\}$. With multiple demonstrations, Dynamic Time Warping (DTW) (Zhang et al., 2024) is applied to temporally align the trajectories.

II) Local Statistics and Heteroscedastic Noise. Instead of assuming homogeneous observation noise, we adapt the noise structure to the variability of the demonstrations. The local mean μ_{in} and standard deviation σ_{in} are computed over y_{in}. A diagonal noise matrix is then defined as $R(t) = \sigma_{\text{in}}^2 I_{y_{\text{in}}}$.

III) Heteroscedastic Initial Kernel. A base kernel (e.g., squared-exponential) K_{base} is initialized using $m(t) = \mu_{\text{in}}$. The initial kernel is constructed by locally scaling the covariance according to the observed variability $K_{\text{init}} = R(t)^2 \odot K_{\text{base}}$, where $\odot$ denotes the Hadamard product. To preserve symmetry and numerical stability, a jitter term ϵI is added, ensuring the matrix remains positive definite and Cholesky-factorizable.

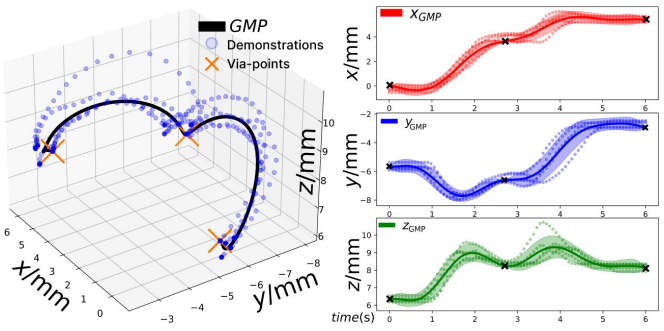

Figure 1: GP-based learning of a 3D spatial trajectory satisfying three via-points. The dimensional breakdown displays the predictive mean alongside its covariance (shaded areas). The strict conditioning of the model is demonstrated by the absence of uncertainty at the via-point locations.

IV) Automatic Kernel Optimization. To obtain the optimal kernel $K_i(t, t)$ without manual tuning, we employ ABC. The optimization problem is formulated as:

$$\pi(K_{\text{init}} \mid t^*, G) \propto f(t^* \mid K_{\text{init}}, G)\,\pi(K_{\text{init}} \mid G) \tag{4}$$

where G represents the model based on the kernel functions, and $\pi(K_{\text{init}} \mid G)$ is the prior distribution grounded in the demonstration data's structure. The ABC algorithm automatically optimizes the hyperparameters, discarding models with a lower probability of convergence.

V) Exact Via-Point Incorporation. A crucial advantage of this formulation is the ability to strictly enforce passage through specific via-points. Let $D_i = \{t, q_i\} \in \mathbb{R}^N$ be the noisy demonstrations, and $\bar{D}_i = \{\bar{t}, \bar{q}_i\} \in \mathbb{R}^{\bar{N}}$ be the set of exact via-points. Defining the augmented sets $\tilde{t} = [t^T, \bar{t}^T]^T$ and $\tilde{q}_i = [q_i^T, \bar{q}_i^T]^T$, the joint distribution is:

$$p\left(\begin{bmatrix} q_i \\ \bar{q}_i \end{bmatrix} \middle\| \begin{bmatrix} t \\ \bar{t} \end{bmatrix}\right) \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{m}(t) \\ \mathbf{m}(\bar{t}) \end{bmatrix}, \begin{bmatrix} K(t,t) + R(t) & K(t,\bar{t}) \\ K(\bar{t},t) & K(\bar{t},\bar{t}) \end{bmatrix}\right) \tag{5}$$

With optimized hyperparameters, the predictive posterior for a point t^* becomes $p(h_i(t^*) \mid t^*, \bar{D}_i) = \mathcal{N}(\tilde{\mu}_i(t^*), \tilde{\Sigma}_i(t^*))$, where:

$$\begin{cases} \tilde{\mu}_i(t^*) = \mathbf{m}(t^*) + K_i(t^*, \tilde{t})K_i'(\tilde{t}, \tilde{t})^{-1}(\tilde{q}_i - \mathbf{m}(\tilde{t})) \\ \tilde{\Sigma}_i(t^*) = K_i(t^*, t^*) - K_i(t^*, t)K_i'(\tilde{t}, \tilde{t})^{-1}K_i(\tilde{t}, t^*) \end{cases} \tag{6}$$

This guarantees that the learned primitive passes exactly through the predefined via-points.

4.2. Real-Time Adaptation via Pathwise Conditioning

While the previous formulation encapsulates the nominal policy, collaborative tasks require continuous, real-time adaptation to changing via-points (e.g., a human moving or task-parameters changing during the execution). Standard GP regression requires recalculating complex matrix inverses upon every observation change, which is computationally prohibitive at high control frequencies (e.g., 30Hz).

To solve this, we adopt a Pathwise Conditioning strategy (Wilson et al., 2021), applying a linear Matheron-style update directly to the sampled prior trajectory rather than the global model. Starting from a prior trajectory $h_i^{\text{prior}}(\cdot)$, the pathwise-conditioned trajectory evaluating the new constraints $\tilde{q}_i$ at times $\tilde{t}$ is defined as:

$$h_i^{\text{post}}(t) = h_i^{\text{prior}}(t) + K_i(t, \tilde{t})\,K_i'(\tilde{t}, \tilde{t})^{-1}(\tilde{q}_i - h_i^{\text{prior}}(\tilde{t})) \tag{7}$$

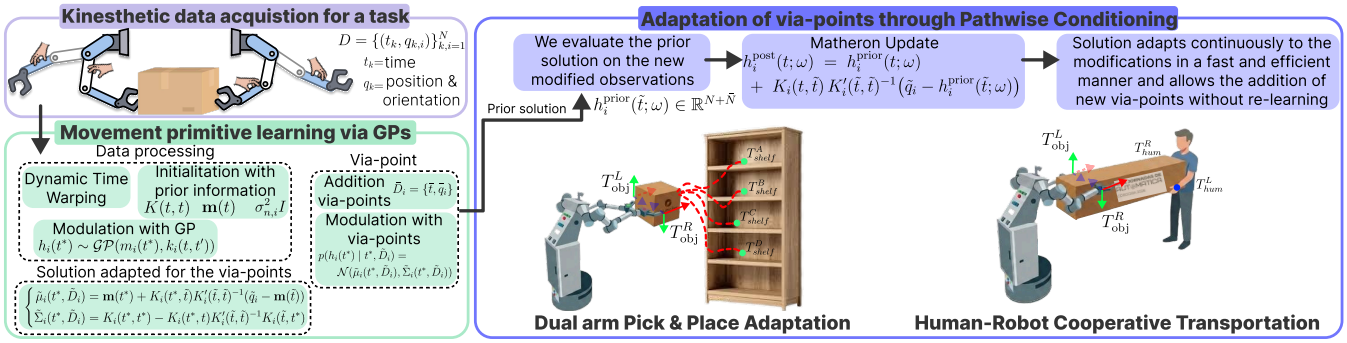

Figure 2: General scheme. First, task demonstrations are acquired. Next, movement primitives are learned using Gaussian Processes: a heteroscedastic initialization of the mean and noise is computed from the demonstrations' envelope, an initial kernel is constructed, and the hyperparameters are automatically optimized. To this learned policy, a linear correction is applied via Pathwise Conditioning to enforce new via-points without retraining, locally preserving the learned primitive and its predictive uncertainty. Website project and videos: `https://adrianprados.github.io/WebPageJornadasAutomatica2026/`

Eq. (7) modifies the prior trajectory via a data-dependent linear correction. We employ independent kernels per output coordinate, decoupling the update by dimension d:

$$\Delta f_d(t) = K^{(d)}(t,\tilde{t})(K^{(d)}(\tilde{t},\tilde{t}) + \Sigma^{(d)}_{\text{obs}})^{-1}(\tilde{q}_{i,d} - h_d^{\text{prior}}(\tilde{t})) \tag{8}$$

yielding the final corrected trajectory $h_d^{\text{post}}(t) = h_d^{\text{prior}}(t) + \Delta f_d(t)$. Crucially, the linear correction is weighted by the cross-covariance $K_i(t,\tilde{t})$. For times t remote from the updated constraint $\tilde{t}$, the correction naturally tends to zero. This preserves the structural intent of the learned primitive while strictly satisfying the new via-points. Furthermore, because this formulation scales across multiple dimensions, it efficiently handles full 6D spatial constraints, adapting both position and orientation simultaneously in a zero-shot manner.

4.3. Task-Parameterized Dual-Arm Coordination

Efficiency and adaptability of the Pathwise Update presented in Eq. (7) provide the mathematical foundation required for multi-agent coordination. In collaborative manipulation, whether dual-arm or human-robot, the agents are physically coupled by the manipulated object, forming a closed kinematic chain. Consequently, the via-points $\tilde{q}_i$ are no longer static spatial targets, but dynamic, coupled constraints dictated by the state of the other agent. Rather than relying on a rigid leader-follower paradigm, our framework generates these relative via-points bidirectionally. The formulation of these relative via-points and the closed-chain synchronization algorithm are detailed in the following section (overview in Fig. 2).

Let $Agent_1$ be the leading agent (e.g., a human operator or the primary arm responding to a dynamic target) and $Agent_2$ be the follower arm. Let $\tilde{q}_1 \in SE(3)$ represent the 6D pose of the leader at a specific update time $\tilde{t}$. The shared payload imposes a strict geometric constraint. We define this task-specific constraint as a relative rigid transformation $T_{\text{obj}} \in SE(3)$, which encapsulates the dimensions and grasp points of the specific object being manipulated (e.g., the width of a box or the length of a table). The mandatory via-point for the follower agent, $\tilde{q}_2$, is calculated instantaneously by composing the leader's pose with the task parameter:

$$\tilde{q}_2(\tilde{t}) = \tilde{q}_1(\tilde{t}) \oplus T_{\text{obj}} \tag{9}$$

where $\oplus$ denotes the composition operator in $SE(3)$. For pure positional tasks in $\mathbb{R}^3$, this simplifies to $\tilde{q}_2 = \tilde{q}_1 + R(\theta_1)\mathbf{d}_{\text{obj}}$, with $\mathbf{d}_{\text{obj}}$ being the task-dependent offset vector and $R(\theta_1)$ the rotation matrix of the leader.

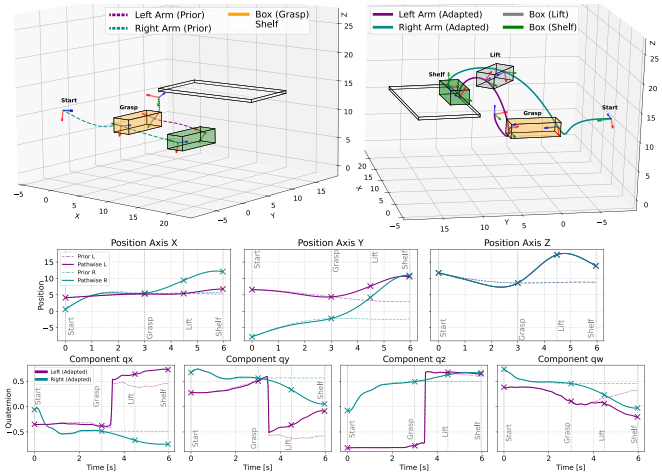

Figure 3: Pathwise update for 6D dual-arm box transport. **(Top)** 3D prior trajectories (left) and simultaneous adaptation to new target POSEs (right). **(Middle)** Prior (dashed) vs. adapted (solid) spatial positions for the left (purple) and right (green) arms. **(Bottom)** Quaternion trajectories. Dynamic via-points (X) across task phases enforce closed-chain constraints.

This kinematic coupling bridges the gap between the physical task and the probabilistic model. As the leader moves or adapts to a new destination, Eq. (9) calculates the exact spatial constraint required to prevent object dropping or motor stress. This dynamically generated $\tilde{q}_2$ is directly injected as the new observation into Eq. (7) for $Agent_2$:

$$h_2^{\text{post}}(t) = h_2^{\text{prior}}(t) + K_2(t,\tilde{t})\, K_2'(\tilde{t},\tilde{t})^{-1}(\tilde{q}_2 - h_2^{\text{prior}}(\tilde{t})) \tag{10}$$

In GP regression, $K_2'(\tilde{t},\tilde{t})$ is evaluated strictly on the temporal inputs ($\tilde{t}$) of the via-points, independent of their spatial targets. Assuming the temporal phase of the constraints remains constant, the most expensive operation—the inversion $K_2'(\tilde{t},\tilde{t})^{-1}$—can be precomputed offline. As the task parameter T_{obj} changes for different objects, or as $\tilde{q}_1$ continuously shifts the spatial and rotational targets in real-time, the algorithm only requires fast matrix-vector multiplications. This allows the follower to adapt its entire trajectory continuously at high frequencies in a zero-shot manner, maintaining closed-chain coordination without any model retraining (see Fig. 3).

5. Experiments and Results

5.1. Task generalization methods comparison in 2D

We compared our method against four task-parameterized models: TP-GPR Calinon (2016), TP-GMM Calinon and Caldwell (2014), TP-ProMP Paraschos et al. (2013), and Synthetic-TP-GMM Prados et al. (2024a). The test scenario relies on a single demonstration where both the starting and ending constraints are dynamically modified. To comprehensively assess the algorithms' adaptability, 4 distinct configurations were evaluated: *final restriction near*, *final restriction far*, *both restrictions near*, and *both restrictions far*. Fig 4 presents the results for the *both restrictions near* case.

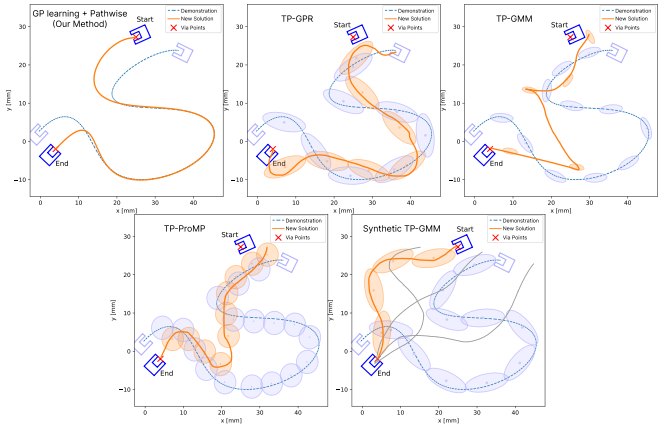

Figure 4: Blue and orange paths represent the original demonstration and the newly generated motion policy, respectively, while the probabilistic variance is captured by the ellipses. While the compared algorithms fail to fulfill the boundary requirements from a single demonstration without generating synthetic data, our method successfully satisfies the geometric constraints.

We assess generalization capabilities using three metrics (Li and Figueroa, 2023), Eq. (11): *Start and Goal Cosine Similarities* ($\cos(\theta_s)$ & $\cos(\theta_g)$), which measure the angular alignment between the generated trajectory vectors (v_s, v_g) and the required target orientations (v_o, v_e), where values approaching 1 denote optimal alignment; and *Endpoint Distance* ($\mathcal{D}$), which quantifies the accumulated spatial error at both boundaries (P_s, P_e) relative to the trajectory endpoints (ξ_s, ξ_g).

$$\cos(\theta_s) = \frac{v_s \cdot v_o}{\|v_s\| \, \|v_o\|} \;,\; \cos(\theta_g) = \frac{v_g \cdot v_e}{\|v_g\| \, \|v_e\|} \;,\; \mathcal{D} = d(\xi_s, P_s) + d(\xi_g, P_e) \tag{11}$$

Compared algorithms (Table 1) struggle to generalize reliably from a single demonstration. Specifically, TP-GPR initiates with an inverted orientation ($\cos(\theta_s) < 0$) and exhibits massive endpoint distance errors. Similarly, TP-ProMP shows significant spatial deviations right from the starting position and fails to produce consistent trajectories. Synthetic-TP-GMM successfully satisfies the geometric constraints by augmenting the dataset with artificial demonstrations. However, this process dilutes the influence of the original motion structure and requires over 2 seconds to compute, rendering it unviable for real-time corrections. In contrast, our method strictly satisfies position and orientation constraints with near-zero error and generates the adapted policy in 0.025 seconds, demonstrating its superior suitability for highly dynamic environments.

Table 1: Comparison of methods over 100 environments ($\mu \pm \sigma$)

Method	$\cos(\theta_s)$	$\cos(\theta_g)$	$\mathcal{D}$	Time(s)
Ours	**0.9999 ± 0.0001**	**0.9998 ± 0.0002**	**0.0002 ± 0.0001**	**0.0249 ± 0.0015**
TP-GPR	-0.9971 ± 0.0150	0.7015 ± 0.1205	0.3987 ± 0.0845	0.5148 ± 0.0420
TP-GMM	0.8015 ± 0.0950	0.5000 ± 0.1500	0.6997 ± 0.1120	0.0402 ± 0.0050
TP-ProMP	0.1742 ± 0.2010	0.9012 ± 0.0540	0.4069 ± 0.0920	0.0401 ± 0.0045
Syn. TP-GMM	0.9498 ± 0.0210	0.9557 ± 0.0185	0.0006 ± 0.0003	2.6912 ± 0.1500

5.2. Robot Experiments: Box Pick-and-Place

Our method is evaluated on a dual-arm pick-and-place task within a shelf environment. The stringent spatial constraints, combined with the requirement to maintain a closed kinematic chain, provide an ideal benchmark for assessing skill generalization. This setup emphasizes the preservation of fundamental human manipulation strategies—including coordinated lifting, stable transport, and high-precision insertion—while demonstrating robust adaptation to novel geometric configurations. The experiment evaluates the framework's capacity to generalize from a single-context demonstration to various previously unseen task settings as in Fig. 5. The baseline dual-arm behavior is acquired through a single kinesthetic demonstration. The manipulation task is inherently partitioned into sequential phases: (i) approach and grasp, (ii) lifting and transport from the initial table to the shelf, and (iii) object placement and release. Leveraging this baseline, the system extrapolates the trajectory to new 6D goal poses within the shelf, ensuring convergence at task boundaries while preserving the intrinsic characteristics of the demonstrated skill.

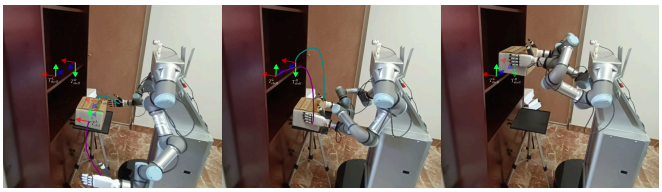

Figure 5: Dual-arm pick-and-place task with ADAM. **(Left)** Approach and grasp, highlighting initial frames (T^L_{obj}, T^R_{obj}) and updated 6D shelf goals $(T^L_{shelf}, T^R_{shelf})$. **(Middle)** Coordinated transport via zero-shot adapted trajectories (purple/teal) seamlessly connecting start and goal poses while maintaining the closed kinematic chain. **(Right)** High-precision placement at novel configurations, proving robust generalization from a single demonstration.

To validate the practical feasibility of this zero-shot adaptation, the generated trajectories were deployed in the ADAM mobile manipulator (Mora et al., 2024), evaluated both in the PyBullet-based ADAMSim environment (Prados et al., 2025a) and under real-world conditions. Throughout the execution, updated 6D target poses were provided. The proposed algorithm generates coordinated trajectories for both arms at a continuous rate of 30 Hz. This ensures that the rigid relative constraints imposed by the payload are strictly maintained across all motion phases, enabling real-time reactivity without the need for model retraining. Videos presented in `https://adrianprados.github.io/WebPageJornadasAutomatica2026/`

5.3. Robot Experiments: Human-Robot Transportation

Our algorithm has been evaluated for large object transportation tasks between a robot and a human (Fig. 6). The initial demonstrations consist solely of approaching the box (T^L_{obj}, T^R_{obj}). The algorithm updates the via-points based on

the keypoints of the user controlling the other side of the object being transported. In this case, using a Yolo Pose-based vision system, the characteristic features (the user's hands) are extracted, thereby determining the positions of both via-points for real-time optimization (T^L_{hum}, T^R_{hum}). The human acts as the leader, while the robot acts as the follower. This enables real-time control, allowing for the coordinated human-robot transportation of large objects. This process has been evaluated both in simulation using ADAMSim and with the real robot. In the real-world task, certain limitations were encountered when generating large orientation modifications. These arise from the robot's physical safety limits as well as the absence of a force control system to assist in updating the via-points. Despite this, our algorithm presents an initial approach to rapidly adapt the coordinated movement between a human and a robot for the collaborative transportation of large items.

Figure 6: Human-robot transport with ADAM. **(Left and Middle)** Real-world execution. The visual tracking system monitors the human operator's hands (T^L_{hum}, T^R_{hum}, marked with blue and orange dots) to continuously compute the via-points for the robot's end-effectors (T^L_{obj}, T^R_{obj}), preserving the closed kinematic chain. **(Right)** Real-time example using ADAMSim.

6. Conclusions and Future Works

This paper presents a method based on Gaussian Processes and Pathwise Conditioning that enables real-time adaptation for collaborative manipulation tasks. The algorithm achieves zero-shot adaptation to new 6D spatial constraints without the need for retraining, ensuring the maintenance of the closed kinematic chain with near-zero error, as demonstrated by the experiments with the ADAM robot. As future work, we propose extending this framework to integrate dynamic obstacle avoidance into the trajectory via Signed Distance Fields. Additionally, the management of interaction forces during collaborative transportation will be investigated to improve safety and fluency in both dual-arm and human-robot tasks.

Acknowledgments

Work supported by Advanced Mobile dual-arm manipulator for Elderly People Attendance (AMME) (PID2022-139227OB-I00) by Ministerio de Ciencia e Innovacion and by iRoboCity2030-CM (TEC-2024/TEC-62) by Programas de Actividades I+D en tecnologías de la Comunidad de Madrid.

References

Arduengo, M., Colomé, A., Borras, J., Torras, C., 2021. Task-adaptive robot learning from demonstration with gaussian process models under replication. IEEE RAL 6 (2), 966–973.

Barber, R., Ortiz, F. J., Garrido, S., Mora, A., Prados, A., Méndez, I., Mozos, Ó. M., 2022. A multirobot system in an assisted home environment to support the elderly in their daily lives. Sensors 22 (20), 7983.

Calinon, S., 2016. A tutorial on task-parameterized movement learning and retrieval. Intelligent service robotics 9 (1), 1–29.

Calinon, S., Caldwell, D. G., 2014. A task-parameterized probabilistic model with minimal intervention control. In: ICRA. IEEE, pp. 39–44.

Carrasco, A. P., Velasco, A. M., García, A. M., Bullón, S. G., Castaño, R. I. B., 2024. Learning from demonstration through synthetic data for parameterized tasks. Jornadas de Automática (45).

Feng, C., Liu, Z., Li, W., Lu, X., Jing, Y., Ma, Y., 2025. Improved gaussian mixture model and gaussian mixture regression for learning from demonstration based on gaussian noise scattering. Adv. Eng. Infor. 65, 103192.

Huang, K., Ji, X., Su, J., Qu, X., 2025. Task-parameterized dynamic movement primitives with reinforcement learning for improved motion planning. IEEE Robotics and Automation Letters.

Lee, Q. Y., Kulkarni, S. R., Yang, L., Noronha, B., Wee, Y., Campolo, D., 2025. Generalizing robot trajectories from single-context human demonstrations: A probabilistic approach. arXiv preprint arXiv:2503.05619.

Li, T., Figueroa, N., 2023. Task generalization with stability guarantees via elastic dynamical system motion policies. In: CORL.

Mehta, S. A., Ciftci, Y. U., Ramachandran, B., Bansal, S., Losey, D. P., 2025. Stable-bc: Controlling covariate shift with stable behavior cloning. IEEE Robotics and Automation Letters.

Mendez, A., Prados, A., Menendez, E., Barber, R., 2024. Everyday objects rearrangement in a human-like manner via robotic imagination and learning from demonstration. IEEE Access 12, 98–119.

Mora, A., Prados, A., Mendez, A., Espinoza, G., Gonzalez, P., Lopez, B., Muñoz, V., Moreno, L., Garrido, S., Barber, R., 2024. Adam: a robotic companion for enhanced quality of life in aging populations. Frontiers in Neurorobotics 18, 1337608.

Paraschos, A., Daniel, C., Peters, J. R., Neumann, G., 2013. Probabilistic movement primitives. Adv. in neural info. processing 26.

Prados, A., Espinoza, G., Mendez, A., Mora, A., Garrido, S., Barber, R., 2025a. Adamsim: Pybullet-based simulation environment for research on domestic mobile manipulator robots. Jornadas de Auto. (46).

Prados, A., Espinoza, G., Moreno, L., Barber, R., 2025b. Coordination of learned decoupled dual-arm tasks through gaussian belief propagation. In: 2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, pp. 15917–15924.

Prados, A., Garrido, S., Barber, R., 2024a. Learning and generalization of task-parameterized skills through few human demonstrations. Eng. Applications of Artificial Intelligence 133, 108310.

Prados, A., Hertel, B., Barber, R., Azadeh, R., 2026a. Elastic fast marching learning from demonstration. Advanced Intelligent Systems 8 (2), e202500607.

Prados, A., Mendez, A., Espinoza, G., Fernandez, N., Barber, R., 2024b. f-divergence optimization for task-parameterized learning from demonstrations algorithm. In: IEEE Int. Conference on Autonomous Robot Systems and Competitions (ICARSC). IEEE, pp. 9–14.

Prados, A., Mora, A., López, B., Muñoz, J., Garrido, S., Barber, R., 2023. Kinesthetic learning based on fast marching square method for manipulation. Applied Sciences 13 (4), 2028.

Prados, A., Moreno, L., Barber, R., 2026b. Learning of movement primitives by gaussian processes from demonstrations. Engineering Applications of Artificial Intelligence 179, 115258.

Qian, K., Xu, X., Liu, H., Bai, J., Luo, S., 2022. Environment-adaptive learning from demonstration for proactive assistance in human–robot collaborative tasks. Robotics and Autonomous Systems 151, 104046.

Ruan, S., Meng, X., Chirikjian, G. S., 2024. Primp: Probabilistically-informed motion primitives for efficient affordance learning from demonstration. IEEE Trans. on Rob. 40, 68–87.

Silvério, J., Rozo, L., Calinon, S., Caldwell, D. G., 2015. Learning bimanual end-effector poses from demonstrations using task-parameterized dynamical systems. In: 2015 IROS. IEEE, pp. 464–470.

Sosa-Ceron, A. D., Gonzalez-Hernandez, H. G., Reyes-Avendaño, J. A., 2022. Learning from demonstrations in human–robot collaborative scenarios: A survey. Robotics 11 (6), 126.

Spencer, J., Choudhury, A., Ziebart, B., Bagnell, J. A., 2021. Feedback in imitation learning: The three regimes of covariate shift. arXiv:2102.02872.

Wilson, J. T., Borovitskiy, V., Terenin, A., Mostowsky, P., Deisenroth, M. P., 2021. Pathwise conditioning of gaussian processes. Journal of Machine Learning Research 22 (105), 1–47.

Zhang, R., Xia, J., Ma, J., Huang, D., Zhang, X., Li, Y., 2024. Human–robot interactive skill learning and correction for polishing based on dynamic time warping iterative learning control. IEEE Transactions on Control Systems Technology.