Supplementary material: Solving Chance Constrained Optimization under Non-Parametric Uncertainty Through Hilbert Space Embedding.

Bharath Gopalakrishnan[1], Arun Kumar Singh[2], K. Madhava Krishna[1] and Dinesh Manocha[3]

I. Extension of the Approach to multiple random variables

A brief version of this extension has been presented in Section V-C of the revised manuscript

It is possible to extend our formulation to multiple random variables, the derivations and the nature of the results would change accordingly. In this rebuttal, we give a detailed derivation that extends the concepts to three random variables $\mathbf{w}_1$, $\mathbf{w}_2$ and $\mathbf{w}_3$. We also demonstrate the following observations:

- The nature of the solution does change when the number of random variables under consideration is increased from two to three. In particular, we observe that incorporating additional random variables leads to conservative solutions in the sense of increase in optimal cost for the same chance constraint probability η.
- We also observe that in spite of increase in conservativeness, the proposed formulation based on RKHS still outperforms existing approaches discussed in the revised manuscript.

Before we get into the derivations and its evaluations, we state certain prerequisites that reflect the the changes in the theory when we incorporate an additional random Variable $\mathbf{w}_3$. All the definitions are taken from the revised manuscript.

Definition of Chance Constraints: The definition of the basic chance constrained optimization (equations 1(a)-1(c)) would now be :

$$\min g(u), \tag{1a}$$

$$p_c(u) \geq \eta, \tag{1b}$$

$$u \in \mathcal{F}, \tag{1c}$$

$$p_c(u) = P(f(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, u) \leq 0) \tag{2}$$

The function $g(u)$ is the user defined cost. $P(.)$ represents probability and $f(.)$ is the constraint function which depends on the decision variable u and uncertain parameters, $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$. The dependence of $f(.)$ on both $\mathbf{w}_1, \mathbf{w}_2$,$\mathbf{w}_3$ and u could possibly be highly non-linear and non-convex. However, we assume $f(.)$ to be separable in u.

Definition Of Kernel Mean Embedding Based on the definitions of Kernel Mean Embedding, (equation (6) of the revised manuscript) now gets modified into:

$$\mu_{p_f}(u) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n} \alpha_i \beta_j \gamma_m k(f(\mathbf{w}_1^i, \mathbf{w}_2^j, \mathbf{w}_3^m, u), .), \tag{3}$$

where, α_i,β_j and γ_m are the weights associated to $\mathbf{w}_1, \mathbf{w}_2$,$\mathbf{w}_3$ computed through reduced set methods.

Reduced set Methods: The revised manuscript has the reduced set described for two random variables $\mathbf{w}_1$ and $\mathbf{w}_2$ (Section II-C). The weights α_i and β_j corresponding to $\mathbf{w}_1$ and $\mathbf{w}_2$ are derived in equations (9)-(10). The extension for the weights of $\mathbf{w}_3$ follows these definitions exactly and can be expressed below as:

$$\gamma_i = \arg\min_{\gamma_m} \| \frac{1}{N} \sum_{i=1}^{i=N} k(\hat{\mathbf{w}}_3{}^i, .) - \sum_{m=1}^{m=n} \gamma_m k(\mathbf{w}_3^m, .) \|_2, s.t \sum \gamma_i = 1, \tag{4}$$

where $\hat{\mathbf{w}}_3^1, \hat{\mathbf{w}}_3^2..\hat{\mathbf{w}}_3^N$ represent N i.i.d samples of $\mathbf{w}_3$ and $\mathbf{w}_3^1..\mathbf{w}_3^n$ represent a reduced set of n samples. It is again implied that $n << N$.

Algebraic form of Constraint Function and its corresponding Embedding: The constraint functions in the revised manuscript (equation (11)) is now modified to:

$$f(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, u) = \sum_{i=0}^{l} h_i(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3)u^i \tag{5}$$

where, $h_i(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3), \Re^n \rightarrow R$ is a generic possibly non-linear function of $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$, while u^i represents a monomial of order i.The kernel mean embedding of $f(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, u)$, would now take the form:

$$\mu_{p_f}(u) = \sum_{i=0}^{i=l} \mu_{h_i} u^i \tag{6}$$

$$\mu_{h_i} = \sum_{i=1}^{i=n} \sum_{j=1}^{j=n} \sum_{m=1}^{m=n} \alpha_i \beta_j \gamma_m k(h_i(\mathbf{w}_1^i, \mathbf{w}_2^j, \mathbf{w}_3^m), .) \tag{7}$$

Construction Of Desired Distribution: Based on all the above definitions, the construction of the desired distribution would change accordingly. The optimization problem used to solve for u_{nom} (eqn. (16a)-(16c)) can be now modified as:

$$u_{nom} = \arg\min g(u) \tag{8a}$$

$$f(\widetilde{\mathbf{w}}_1^i, \widetilde{\mathbf{w}}_2^j, \widetilde{\mathbf{w}}_3^m, u) \leq 0, \forall i = 1, 2..n_{\mathbf{w}_1}, j = 1, 2..n_{\mathbf{w}_2}, m = 1, 2..n_{\mathbf{w}_3} \tag{8b}$$

$$u \in \mathcal{F} \tag{8c}$$

The kernel mean embedding of the desired distribution can now be computed in the following manner:

$$\mu_{p_f^{des}} = \sum_{i=1}^{i=n_{\mathbf{w}_1}} \sum_{j=1}^{j=n_{\mathbf{w}_2}} \sum_{m=1}^{m=n_{\mathbf{w}_3}} \lambda_i \xi_j \phi_m k(f(\widetilde{\mathbf{w}}_1^i, \widetilde{\mathbf{w}}_2^j, \widetilde{\mathbf{w}}_3^m, u_{nom}), .), \widetilde{\mathbf{w}}_1^i, \widetilde{\mathbf{w}}_2^j, \widetilde{\mathbf{w}}_3^m \in \mathcal{C}_{\widetilde{\mathbf{w}}_1}, \mathcal{C}_{\widetilde{\mathbf{w}}_2}, \mathcal{C}_{\widetilde{\mathbf{w}}_3} \tag{9}$$

Where, λ_i, ξ_j and ϕ_m are constants derived from the reduced set methods described in Section II-C of the revised manuscript.

Main Derivation:

Based on the discussions above we extend the derivations of section III-E as follows to incorporate the random variable $\mathbf{w}_3$: We have

$$\|\mu_{p_f}(u) - \mu_{p_f^{des}}\|^2 = \langle \mu_{p_f}(u) - \mu_{p_f^{des}}, \mu_{p_f}(u) - \mu_{p_f^{des}} \rangle$$

$$= \langle \mu_{h_0} + \mu_{h_1}u + \mu_{h_2}u^2, \mu_{h_0} + \mu_{h_1}u + \mu_{h_2}u^2 \rangle$$

$$-2\langle \mu_{h_0} + \mu_{h_1}u + \mu_{h_2}u^2, \mu_{P_f^{des}} \rangle + \langle \mu_{p_f^{des}}, \mu_{p_f^{des}} \rangle \tag{10}$$

Expanding $\langle \mu_{h_0} + \mu_{h_1}u + \mu_{h_2}u^2, \mu_{h_0} + \mu_{h_1}u + \mu_{h_2}u^2 \rangle$, we get

$$\begin{aligned}
&u^4\langle \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{l=1}^{n} \alpha_i\beta_j\gamma_m k(h_2(\mathbf{w}_1^i, \mathbf{w}_2^j, \mathbf{w}_3^m), .), \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n} \alpha_i\beta_j\gamma_m k(h_2(\mathbf{w}_1^i, \mathbf{w}_2^j, \mathbf{w}_3^m), .)\rangle \\
&+2u^3\langle \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n} \alpha_i\beta_j\gamma_m k(h_2(\mathbf{w}_1^i, \mathbf{w}_2^j, \mathbf{w}_3^m), .), \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n} \alpha_i\beta_j\gamma_m k(h_1(\mathbf{w}_1^i, \mathbf{w}_2^j, \mathbf{w}_3^m), .), .\rangle \\
&+2u^2\langle \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n} \alpha_i\beta_j\gamma_m k(h_2(\mathbf{w}_1^i, \mathbf{w}_2^j, \mathbf{w}_3^m), .), \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n} \alpha_i\beta_j\gamma_m k(h_0(\mathbf{w}_1^i, \mathbf{w}_2^j, \mathbf{w}_3^m), .)\rangle \\
&+u^2\langle \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n} \alpha_i\beta_j\gamma_m k(h_1(\mathbf{w}_1^i, \mathbf{w}_2^j, \mathbf{w}_3^m), .), \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n} \alpha_i\beta_j\gamma_m k(h_1(\mathbf{w}_1^i, \mathbf{w}_2^j, \mathbf{w}_3^m), .)\rangle \\
&+2u\langle \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n} \alpha_i\beta_j\gamma_m k(h_1(\mathbf{w}_1^i, \mathbf{w}_2^j, \mathbf{w}_3^m), .), \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n} \alpha_i\beta_j\gamma_m k(h_0(\mathbf{w}_1^i, \mathbf{w}_2^j, \mathbf{w}_3^m), .)\rangle \\
&+\langle \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n} \alpha_i\beta_j\gamma_m k(h_0(\mathbf{w}_1^i, \mathbf{w}_2^j, \mathbf{w}_3^m), .), \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n} \alpha_i\beta_j\gamma_m k(h_0(\mathbf{w}_1^i, \mathbf{w}_2^j, \mathbf{w}_3^m), .)\rangle
\end{aligned} \tag{11}$$

Using the kernel trick [1] reduces to the following expression

$$\begin{aligned}
&u^4\mathbf{c}_{\alpha\beta\gamma}\mathbf{K}_{h_2h_2}\mathbf{c}_{\alpha\beta\gamma}^T + 2u^3\mathbf{c}_{\alpha\beta\gamma}\mathbf{K}_{h_2h_1}\mathbf{c}_{\alpha\beta\gamma}^T + 2u^2\mathbf{c}_{\alpha\beta\gamma}\mathbf{K}_{h_2h_0}\mathbf{c}_{\alpha\beta\gamma}^T \\
&+u^2\mathbf{c}_{\alpha\beta\gamma}\mathbf{K}_{h_1h_1}\mathbf{c}_{\alpha\beta\gamma}^T + 2u\mathbf{c}_{\alpha\beta\gamma}\mathbf{K}_{h_1h_0}\mathbf{c}_{\alpha\beta\gamma}^T + \mathbf{c}_{\alpha\beta\gamma}\mathbf{K}_{h_0h_0}\mathbf{c}_{\alpha\beta\gamma}^T,
\end{aligned} \tag{12}$$

where,

$$\mathbf{c}_{\alpha\beta\gamma} = [\alpha_1\beta_1\gamma_1, ..., \alpha_1\beta_n\gamma_n, \alpha_2\beta_1\gamma_1, ...\alpha_n\beta_n\gamma_n]_{1X(n*n*n)} \tag{13}$$

$$
\mathbf{K}_{h_i h_j} = \begin{pmatrix}
\mathbf{K}^{111}_{h_i,h_j} & \mathbf{K}^{112}_{h_i,h_j} & \mathbf{K}^{113}_{h_i,h_j} & .. & .. & \mathbf{K}^{11n}_{h_i,h_j} \\
\mathbf{K}^{121}_{h_i,h_j} & \mathbf{K}^{122}_{h_i,h_j} & \mathbf{K}^{123}_{h_i,h_j} & .. & .. & \mathbf{K}^{12n}_{h_i,h_j} \\
\mathbf{K}^{131}_{h_i,h_j} & \mathbf{K}^{132}_{h_i,h_j} & \mathbf{K}^{133}_{h_i,h_j} & .. & .. & \mathbf{K}^{13n}_{h_i,h_j} \\
.. & .. & .. & .. & .. & .. \\
\mathbf{K}^{1n1}_{h_i,h_j} & \mathbf{K}^{1n2}_{h_i,h_j} & \mathbf{K}^{1n3}_{h_i,h_j} & .. & .. & \mathbf{K}^{1nn}_{h_i,h_j} \\
.. & .. & .. & .. & .. & .. \\
\mathbf{K}^{nn1}_{h_i,h_j} & \mathbf{K}^{nn2}_{h_i,h_j} & \mathbf{K}^{nn3}_{h_i,h_j} & .. & .. & \mathbf{K}^{nnn}_{h_i,h_j}
\end{pmatrix} \tag{14}
$$

$$
\mathbf{K}^{abc}_{h_i h_j} =
$$

$$
\begin{pmatrix}
k(h_i(\mathbf{w}_1^a,\mathbf{w}_2^b,\mathbf{w}_3^1),h_j(\mathbf{w}_1^c,\mathbf{w}_2^1,\mathbf{w}_3^1)).. & k(h_i(\mathbf{w}_1^a,\mathbf{w}_2^b,\mathbf{w}_3^1),h_j(\mathbf{w}_1^c,\mathbf{w}_2^1,\mathbf{w}_3^n)).. & k(h_i(\mathbf{w}_1^a,\mathbf{w}_2^b,\mathbf{w}_3^1),h_j(\mathbf{w}_1^c,\mathbf{w}_2^n,\mathbf{w}_3^n)) \\
k(h_i(\mathbf{w}_1^a,\mathbf{w}_2^b,\mathbf{w}_3^2),h_j(\mathbf{w}_1^c,\mathbf{w}_2^1,\mathbf{w}_3^1)).. & k(h_i(\mathbf{w}_1^a,\mathbf{w}_2^b,\mathbf{w}_3^2),h_j(\mathbf{w}_1^c,\mathbf{w}_2^1,\mathbf{w}_3^n)).. & k(h_i(\mathbf{w}_1^a,\mathbf{w}_2^b,\mathbf{w}_3^2),h_j(\mathbf{w}_1^c,\mathbf{w}_2^n,\mathbf{w}_3^n)) \\
.. & .. & .. \\
k(h_i(\mathbf{w}_1^a,\mathbf{w}_2^b,\mathbf{w}_3^n),h_j(\mathbf{w}_1^c,\mathbf{w}_2^1,\mathbf{w}_3^1)).. & k(h_i(\mathbf{w}_1^a,\mathbf{w}_2^b,\mathbf{w}_3^n),h_j(\mathbf{w}_1^c,\mathbf{w}_2^1,\mathbf{w}_3^n)).. & k(h_i(\mathbf{w}_1^a,\mathbf{w}_2^b,\mathbf{w}_3^n),h_j(\mathbf{w}_1^c,\mathbf{w}_2^n,\mathbf{w}_3^n))
\end{pmatrix}_{n\times n^2} \tag{15}
$$

Following a similar process, the second term, $2\langle \mu_{h_0} + \mu_{h_1}u + \mu_{h_2}u^2, \mu_{P_f^{des}} \rangle$ reduces to

$$
2(\mathbf{c}_{\alpha\beta\gamma}\mathbf{K}_{h_2 f}\mathbf{c}^T_{\lambda\xi\phi}u^2 + \mathbf{c}_{\alpha\beta\gamma}\mathbf{K}_{h_1 f}\mathbf{c}^T_{\lambda\xi\phi}u + \mathbf{c}_{\alpha\beta\gamma}\mathbf{K}_{h_0 f}\mathbf{c}^T_{\lambda\xi\phi}) \tag{16}
$$

Where,

$$
\mathbf{c}_{\alpha\beta\gamma} = [\alpha_1\beta_1\gamma_1, ..., \alpha_1\beta_n\gamma_n, \alpha_2\beta_1\gamma_1, ..., \alpha_n\beta_n\gamma_n]_{1X(n*n*n)}
$$

$$
\mathbf{c}_{\lambda\xi\phi} = [\lambda_1\xi_1\phi_1, ..., \lambda_1\xi_n\phi_n, \lambda_2\xi_1\phi_1, ..., \lambda_n\beta_n\gamma_n]_{1X(n_{\mathbf{w}_1}*n_{\mathbf{w}_2}*n_{\mathbf{w}_3})}
$$

$$
\mathbf{K}_{h_i f} = \begin{pmatrix}
\mathbf{K}^{111}_{h_i,f} & \mathbf{K}^{112}_{h_i,f} & \mathbf{K}^{113}_{h_i,f} & .. & .. & \mathbf{K}^{11n_{\mathbf{w}_1}}_{h_i,f} \\
\mathbf{K}^{121}_{h_i,f} & \mathbf{K}^{122}_{h_i,f} & \mathbf{K}^{123}_{h_i,f} & .. & .. & \mathbf{K}^{12n_{\mathbf{w}_1}}_{h_i,f} \\
\mathbf{K}^{131}_{h_i,f} & \mathbf{K}^{132}_{h_i,f} & \mathbf{K}^{133}_{h_i,f} & .. & .. & \mathbf{K}^{13n_{\mathbf{w}_1}}_{h_i,f} \\
.. & .. & .. & .. & .. & .. \\
\mathbf{K}^{1n1}_{h_i,f} & \mathbf{K}^{1n2}_{h_i,h_j} & \mathbf{K}^{1n3}_{h_i,h_j} & .. & .. & \mathbf{K}^{1nn_{\mathbf{w}_1}}_{h_i,h_j} \\
.. & .. & .. & .. & .. & .. \\
\mathbf{K}^{nn1}_{h_i,h_j} & \mathbf{K}^{nn2}_{h_i,h_j} & \mathbf{K}^{nn3}_{h_i,h_j} & .. & .. & \mathbf{K}^{nnn_{\mathbf{w}_1}}_{h_i,h_j}
\end{pmatrix}_{(n^3)\times(n_{\mathbf{w}_1}*n_{\mathbf{w}_2}*n_{\mathbf{w}_3})} \tag{18}
$$

$$\mathbf{K}_{h_i f}^{abc} =$$

$$\begin{pmatrix} k(h_i(\mathbf{w}_1^a, \mathbf{w}_2^b, \mathbf{w}_3^1), f(\tilde{\mathbf{w}}_1^c, \tilde{\mathbf{w}}_2^1, \tilde{\mathbf{w}}_3^1, u_{nom})), .. & k(h_i(\mathbf{w}_1^a, \mathbf{w}_2^b, \mathbf{w}_3^1), f(\tilde{\mathbf{w}}_1^c, \tilde{\mathbf{w}}_2^1, \tilde{\mathbf{w}}_3^{n_{\mathbf{w}3}}, u_{nom}) ... & k(h_i(\mathbf{w}_1^a, \mathbf{w}_2^b, \mathbf{w}_3^2), f(\tilde{\mathbf{w}}_1^c, \tilde{\mathbf{w}}_2^{n_{\mathbf{w}2}}, \tilde{\mathbf{w}}_3^{n_{\mathbf{w}3}}, u_{nom})) \\ k(h_i(\mathbf{w}_1^a, \mathbf{w}_2^b, \mathbf{w}_3^2), f(\tilde{\mathbf{w}}_1^c, \tilde{\mathbf{w}}_2^1, \tilde{\mathbf{w}}_3^1, u_{nom})), .. & k(h_i(\mathbf{w}_1^a, \mathbf{w}_2^b, \mathbf{w}_3^2), f(\tilde{\mathbf{w}}_1^c, \tilde{\mathbf{w}}_2^1, \tilde{\mathbf{w}}_3^{n_{\mathbf{w}3}}, u_{nom})) ... & k(h_i(\mathbf{w}_1^a, \mathbf{w}_2^b, \mathbf{w}_3^1), f(\tilde{\mathbf{w}}_1^c, \tilde{\mathbf{w}}_2^{n_{\mathbf{w}2}}, \tilde{\mathbf{w}}_3^{n_{\mathbf{w}3}}, u_{nom})) \\ \cdots & \cdots & \cdots \\ k(h_i(\mathbf{w}_1^a, \mathbf{w}_2^b, \mathbf{w}_3^n), f(\tilde{\mathbf{w}}_1^c, \tilde{\mathbf{w}}_2^1, \mathbf{w}_3^1, u_{nom})), .. & k(h_i(\mathbf{w}_1^a, \mathbf{w}_2^b, \mathbf{w}_3^n), f(\tilde{\mathbf{w}}_1^c, \tilde{\mathbf{w}}_2^1, \tilde{\mathbf{w}}_3^{n_{\mathbf{w}3}}, u_{nom}) ... & k(h_i(\mathbf{w}_1^a, \mathbf{w}_2^b, \mathbf{w}_3^n), f(\tilde{\mathbf{w}}_1^c, \tilde{\mathbf{w}}_2^{n_{\mathbf{w}2}}, \tilde{\mathbf{w}}_3^{n_{\mathbf{w}3}}, u_{nom})) \end{pmatrix}_{n \times n_{w_2} * n_{w_3}} \tag{19}$$

Finally, the last term, $\langle \mu_{p_f^{des}}, \mu_{p_f^{des}} \rangle$ in (10) can be handled in a similar manner and thus the entire, expression in (10) can be expressed as the following non-linear optimization problem

$$\min \rho_1(a_1 u^4 + a_2 u^3 + a_3 u^2 + a_4 u + a_5) + \rho_2 g(u) \tag{20a}$$

$$u \in \mathcal{F} \tag{20b}$$

$$a_1 = \mathbf{c}_{\alpha\beta\gamma} \mathbf{K}_{h_2 h_2} \mathbf{c}_{\alpha\beta\gamma}^T, a_2 = 2\mathbf{c}_{\alpha\beta\gamma} \mathbf{K}_{h_2 h_1} \mathbf{c}_{\alpha\beta\gamma}^T$$

$$a_3 = 2\mathbf{c}_{\alpha\beta\gamma} \mathbf{K}_{h_2 h_0} \mathbf{c}_{\alpha\beta\gamma}^T + \mathbf{c}_{\alpha\beta\gamma} \mathbf{K}_{h_1 h_1} \mathbf{c}_{\alpha\beta\gamma}^T - 2\mathbf{c}_{\alpha\beta\gamma} \mathbf{K}_{h_2 f} \mathbf{c}_{\lambda\xi\phi}^T$$

$$a_4 = 2\mathbf{c}_{\alpha\beta\gamma} \mathbf{K}_{h_1 h_0} \mathbf{c}_{\alpha\beta\gamma}^T - 2\mathbf{c}_{\alpha\beta\gamma} \mathbf{K}_{h_1 f} \mathbf{c}_{\lambda\xi\phi}^T$$

$$a_5 = \mathbf{c}_{\alpha\beta\gamma} \mathbf{K}_{h_0 h_0} \mathbf{c}_{\alpha\beta\gamma}^T - 2\mathbf{c}_{\alpha\beta\gamma} \mathbf{K}_{h_0 f} \mathbf{c}_{\lambda\xi\phi}^T + \mathbf{c}_{\lambda\xi\phi} \mathbf{K}_{ff} \mathbf{c}_{\lambda\xi\phi}^T$$

Numerical Evaluation:

We evaluate the effect of incorporating an additional random variable $\mathbf{w}_3$ for the example of the *inverse dynamics based path tracking for a 2-link manipulator*. We modify the definitions of the constrained optimization problem presented in equations 39(a)-39(c) of the revised manuscript to accommodate $\mathbf{w}_3$, as shown below:

$$\arg\min_{\ddot{\mathbf{q}}(t)} \frac{1}{2} \|\mathbf{J}(\overline{\mathbf{q}}(t))\ddot{\mathbf{q}}(t) + \dot{\mathbf{J}}(\overline{\mathbf{q}}(t), \dot{\overline{\mathbf{q}}}(t))\dot{\mathbf{q}}(t) - \ddot{\mathbf{x}}(t)\|_2^2 \tag{22a}$$

$$P(f_i(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, u_1, u_2) \leq 0) \geq \eta \tag{22b}$$

$$|\ddot{\mathbf{q}}(t)| \leq \ddot{\mathbf{q}}_{max} \tag{22c}$$

The constraint $f_i(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, u_1, u_2)$ are given below:

$$f_i(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, u_1, u_2) = \sum_{j=1}^{j=2} h_i^j(\mathbf{w}_1.\mathbf{w}_2, \mathbf{w}_3)u_j(t) + h_i(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3) \leq 0 \quad \forall i = 1, 2 \tag{23}$$

$$\text{where, } \mathbf{w}_1 = (q_1(t), q_2(t)), \mathbf{w}_2 = (\dot{q}_1(t), \dot{q}_2(t)), \tag{24}$$

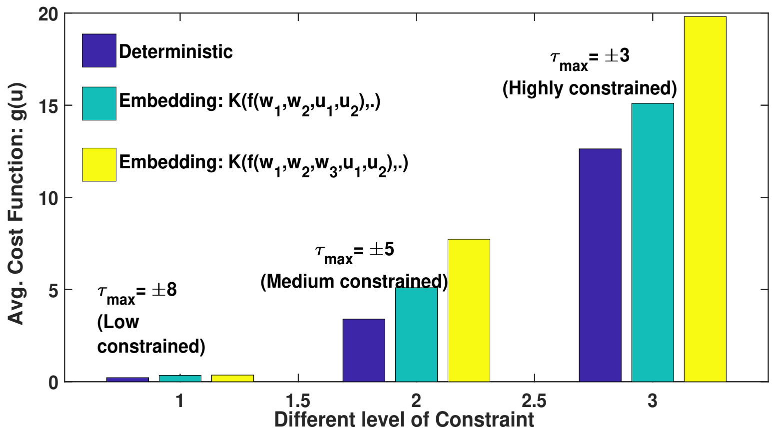
$$\mathbf{w}_3 = (l_1, l_2, m_1, m_2, r_1, r_2, Iz_1, Iz_2)$$

$$(u_1, u_2) = (\ddot{q}_1(t), \ddot{q}_2(t))$$

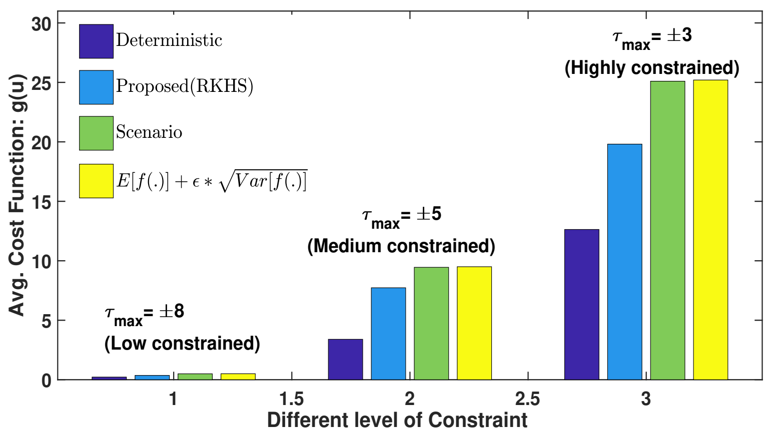
Here $\mathbf{w}_3$ is the collection of all the variables related to the physical dimension of the manipulator i.e, l_i, m_i and r_i, represent the length, mass and the width of the i^{th} link of the manipulator respectively and Iz_i represents the moment of inertia of the i^{th} link of the manipulator.

The results are summarized in the Fig.1(a)-1(b). In Fig.1(a), we show the increase in the optimal cost upon inclusion of more uncertainty in the problem through $\mathbf{w}_3$. Fig.1(b) shows that even with increase with uncertainty, our RKHS embedding based formulation outperforms existing methods like scenario approximation and approaches based on surrogate constraints (3) of the revised manuscript.

[1] B. Schölkopf, K. Muandet, K. Fukumizu, S. Harmeling, and J. Peters, "Computing functions of random variables via reproducing kernel hilbert space representations," *Statistics and Computing*, vol. 25, no. 4, pp. 755–766, 2015.



(a)



(b)

Fig. 1 Figure 1(a):Shows the effect of incorporating w_3 into the proposed RKHS formulation. It is clearly observed that as the constraints become stiffle, the average cost incurred is high when w_3 is incorporated into the formulation. Figure 1(b) shows that inspite of the proposed solution being conservative under the influence of w_3, it still performs better than the other methods like the scenario approach and the expectation - variance method mentioned in equation (3) of the main paper.