We would like to thank the reviewers for their useful feedback. We first address the comment shared by all the reviewers concerning the practical applications of our work. Then, we address other comments individually. Notice that the extra page of content allowed in the final version will allow us to improve the paper according to your recommendations.

Practical applications of HD-GaBO (answer to all the reviewers) First of all, we would like to emphasize that our manuscript is mainly a theoretical contribution and so it aims at introducing the theoretical tools for the proposed HD-GaBO framework. However, we agree with the reviewers that presenting and discussing potential applications may make our motivation clearer and led to future applied extensions of our approach. Therefore we propose to add the following paragraph in the introduction of the paper. Our latent Riemannian manifolds may be exploited for the optimization of controller parameters in robotics. Of particular interest is the optimization of the error gain matrix $\boldsymbol{Q}_t \in \mathcal{S}_{++}^{D_x}$ and control gain matrix $\boldsymbol{R}_t \in \mathcal{S}_{++}^{D_u}$ in linear quadratic regulators (LQR), where D_x and D_u are the dimensionality of the system state and control input, respectively. The system state may consist of the linear and angular position and velocity of the robot end-effector, so $D_x = 13$, and D_u corresponds to Cartesian accelerations or wrench commands. Along some parts of the robot trajectory, the error w.r.t. some dimensions of the state space may not influence the execution of the task (i.e., insignificant effects on the LQR cost function), so that the matrix $\boldsymbol{Q}_t$ for this trajectory segment may be efficiently optimized in a latent space $\mathcal{S}_{++}^{d_x}$ with $d_x < D_x$. A similar analysis applies for $\boldsymbol{R}$. Notice that, although BO has been applied to optimize LQR parameters *[A. Marco et al. Automatic LQR tuning based on Gaussian process global optimization. ICRA, 2016.]*, the problem was greatly simplified as only diagonal matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$ were considered in the optimization, resulting in a loss of flexibility in the controller. From a broader point of view, the low-dimensional assumption may also apply in the optimization of gain matrices for other types of controllers. Another interesting application is the identification of dynamic model parameters of (highly-)redundant robots. These parameters typically include the inertia matrix $\boldsymbol{M} \in \mathcal{S}_{++}^{D}$ with D being the number of robot joints. As discussed in *[S. Zhu et al. Efficient model identification for tensegrity locomotion. IROS, 2018.]*, a low-dimensional representation of the parameter space and state-action space may be sufficient to determine the system dynamics. Therefore, the inertia matrix may be more efficiently represented in a lower-dimensional SPD latent space. A third application concerns the optimization of object shape representations. Indeed, shape spaces are typically characterized on high-dimensional unit spheres $\mathcal{S}^D$. Several works have shown that the main features of the shapes are efficiently represented in a low-dimensional latent space $\mathcal{S}^d$ inheriting the geometry of the original manifold (see e.g.,*[S. Jung et al. Analysis of principal nested spheres. Biometrika, 99(3):551–568, 2012.]*). Therefore, such latent spaces may be exploited for shape representation optimization.

Answers to Reviewer 1 We appreciate the positive feedback of the reviewer. We will correct the indicated typos and discuss the suggested related works. We agree that the inverse map does not necessarily exist if the manifold contains self-intersection. In this case, a possibility would be to learn a non-parametric reconstruction mapping minimizing the sum of the squared residuals, e.g., based on a wrapped Gaussian process *[A. Mallasto et al. Wrapped Gaussian Process Regression on Riemannian Manifolds. CVPR, 2018.]*. However, most of the Riemannian manifolds encountered in machine learning and robotics applications do not self-intersect, so that this problem is avoided.

Answers to Reviewer 2 We agree that prior knowledge on the geometric structure is necessary. However, an advantage of HD-GaBO is its ability to exploit this prior knowledge to build a latent manifold from nested reconstructions that take advantage of the known geometry. We will add an algorithmic summary at the beginning of Section 3 and also discuss the suggested related works, specially that of D. Duvenaud on topological manifolds. Model dimensionality mismatch is a common problem in many dimensionality reduction methods. Intuitively, if d is unknown and estimated lower than the real d, the optimum of the function may not be included in the estimated latent space, leading to a suboptimal solution. In order to attenuate this effect, we hypothesize that the dimension d should be selected slightly higher in case of uncertainty on its value. We will investigate this point more thoroughly in our future work.

Answers to Reviewer 3 The examples presented in Fig. 1 aim at illustrating our assumption, as high-dimensional parameter spaces are hard to visualize. However, these examples are extensible to higher dimensions. Concerning the comment about what we mean with an objective function that only varies within a low-dimensional latent space, this implies that some dimensions of the parameter space do not influence the value of the function. By structure-preserving mapping, we understand a projection from a high-dimensional manifold to a low-dimensional manifold which has the same geometric properties in the lower-dimensional space. We will precise these points in the paper. Moreover, we indeed assume that the map m^{-1} is a right inverse. We will rename it for sake of clarity.

Answers to Reviewer 4 We believe that the contribution of our paper goes beyond the combination of the ideas presented in [26] and [35]. Indeed, the use of nested mappings is an important theoretical component of our framework and provides an interpretable dimensionality reduction strategy in contrast to random embeddings. Moreover, although the idea of using Euclidean latent spaces in BO is quite popular, to the best of our knowledge, latent spaces that preserve the geometric properties of a non-Euclidean spaces have not been investigated in BO. As stated previously, our manuscript is mainly a theoretical contribution. However, we evaluated our framework on classical benchmark functions widely used in BO, which we projected on Riemannian manifolds. We consider the application of HD-GaBO on complex real experiments (see applications above) as part of a future journal work (due to space constraints).