Chemically-informed data-driven optimization (ChIDDO): Leveraging physical models and Bayesian learning to accelerate chemical research

(Electronic Supplementary Information)

Daniel Frey, Juhee Shin, Christopher Musco, and Miguel A. Modestino

Acquisition Function Calculations

Expected Improvement (EI):

$$EI(x) = \left(\mu(x) - f\left(x^+\right) - \xi\right)\psi\left(\frac{\mu(x) - f\left(x^+\right) - \xi}{\sigma(x)}\right) + \sigma(x)\phi\left(\frac{\mu(x) - f\left(x^+\right) - \xi}{\sigma(x)}\right)$$

where $\mu(x)$ is the mean of the regressor at x, $\sigma(x)$ is the variance of the regressor at x, f is the function to be maximized, x^+ is the location of the estimated maximum, ξ is the exploration/exploitation parameter, and $\psi(z)$ is the cumulative distribution function of a standard Guassian distribution, and $\phi(z)$ is the density function of a standard Gaussian distribution[1].

Probability of Improvement (PI):

$$PI(x) = \psi\left(\frac{\left(\mu(x) - f\left(x^+\right) - \xi\right)}{\sigma(x)}\right)$$

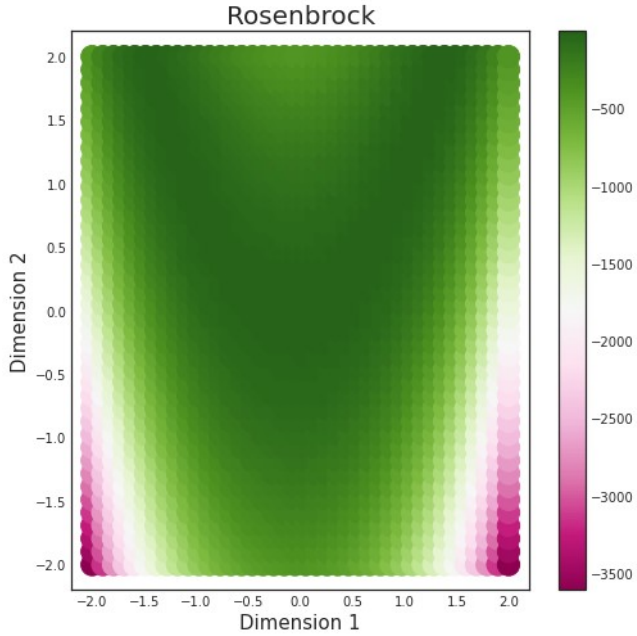
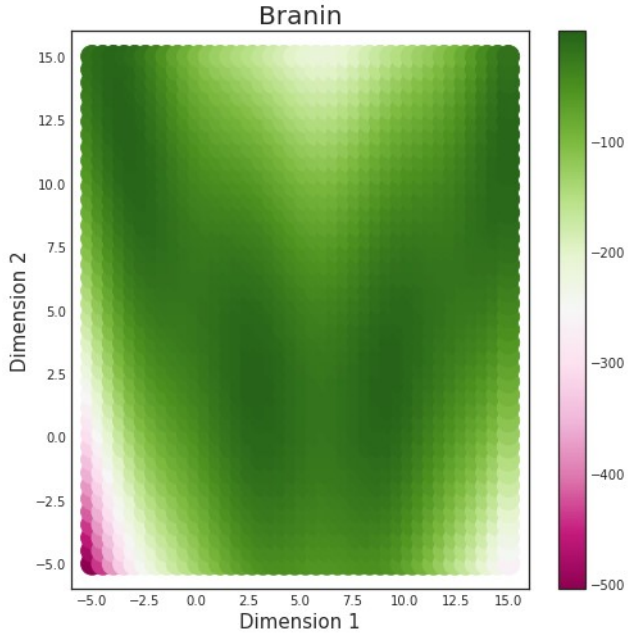
where $\mu(x)$ is the mean of the regressor at x, $\sigma(x)$ is the variance of the regressor at x, f is the function to be maximized, x^+ is the location of the estimated maximum, ξ is the exploration/exploitation parameter, and $\psi(z)$ is the cumulative distribution function of a standard Guassian distribution[1].

Upper Confidence Bound (UCB):

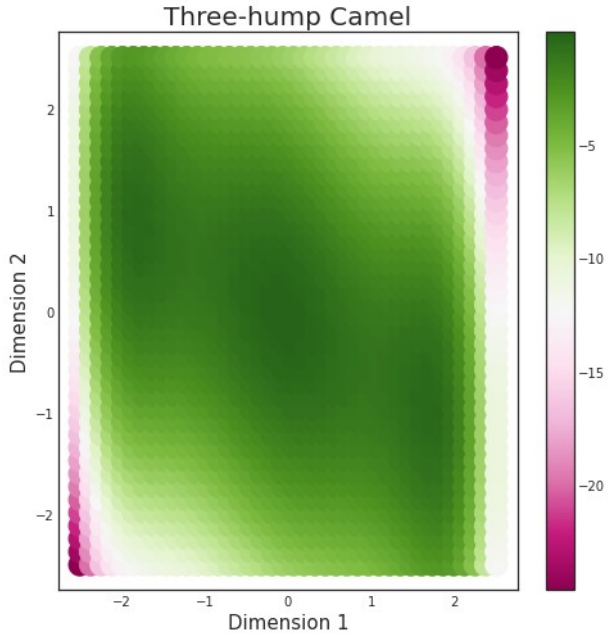
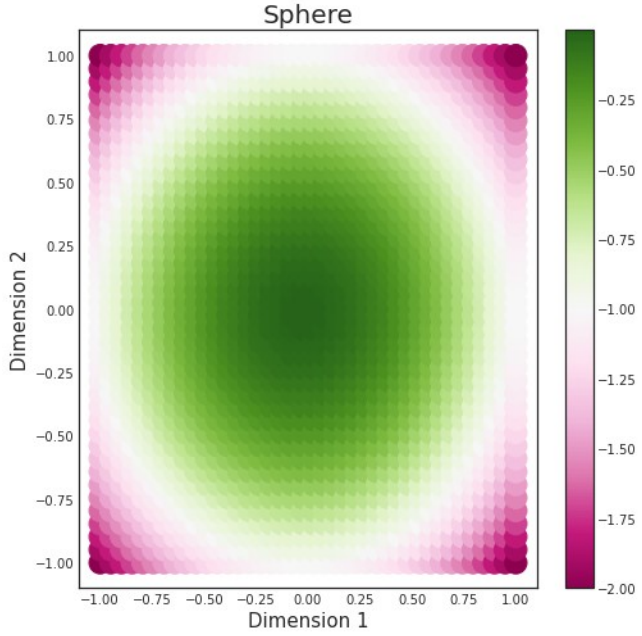
$$UCB(x) = \mu(x) + \beta\sigma(x)$$

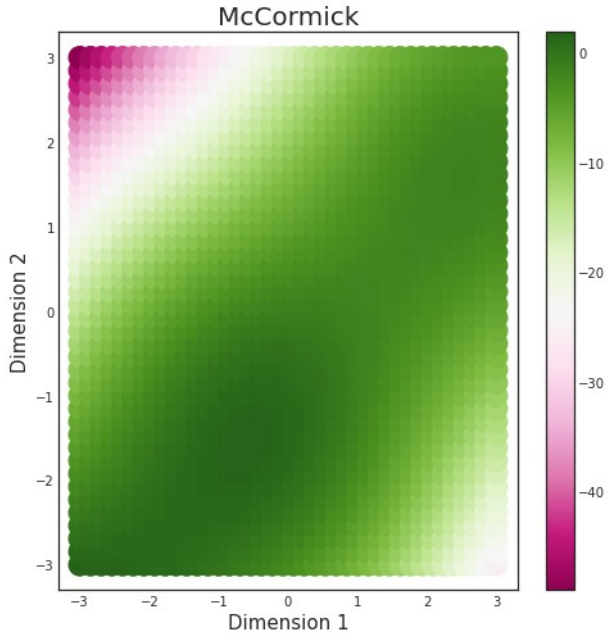
where $\mu(x)$ is the mean of the regressor at x, $\sigma(x)$ is the variance of the regressor at x, and β is the exploration/exploitation parameter[1].

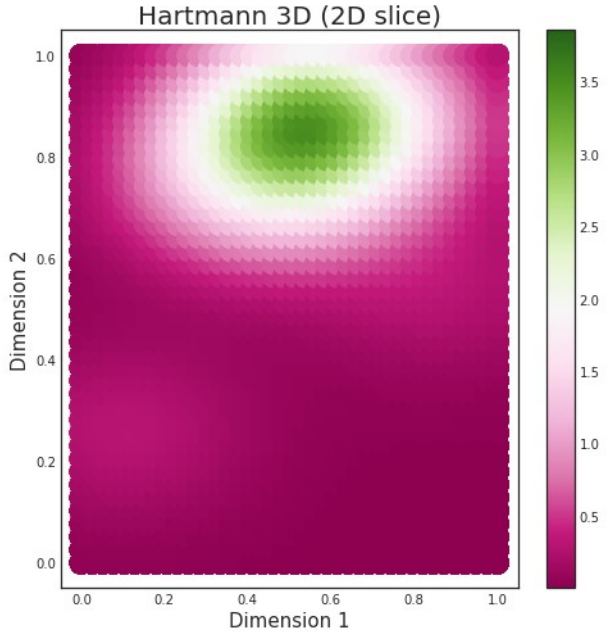
Objective Function Descriptions

<table>
<tr><td> 
</td><td rowspan="2">$P_1 = 100 \pm 20$
$P_2 = 1 \pm 50$
$P_3 = 0 \pm 1$
$P_4 = 0 \pm 1$</td></tr>
<tr><td>$$f(x) = \sum_{i=1}^{d-1} \left[P_1 \left((x_{i+1} + P_3) - x_i^2 \right)^2 + \left(P_2 - (x_i + P_4) \right)^2 \right]$$</td></tr>
<tr><td></td><td></td></tr>
<tr><td> 
</td><td rowspan="2">$a = P_1$
$b = P_2/(4\pi^2)$
$c = P_3/\pi$
$r = P_4$
$s = P_5$
$t = P_6/(8\pi)$

$P_1 = 1 \pm 0.9$
$P_2 = 5.1 \pm 2$
$P_3 = 5 \pm 2$
$P_4 = 6 \pm 4$
$P_5 = 10 \pm 5$
$P_6 = 1 \pm 0.5$</td></tr>
<tr><td>$$f(x) = a\left(x_2 - bx_1^2 + cx_1 - r\right)^2 + s(1-t)\cos\left(x_1\right) + s$$</td></tr>
<tr><td></td><td></td></tr>
</table>

	$P_1 = 2 \pm 1$ $P_2 = 1.05 \pm 0.2$ $P_3 = 1 \pm 0.3$ $P_4 = 1 \pm 0.8$ $P_5 = 1 \pm 0.7$
$f(x) = P_1 x_1^2 - P_2 x_1^4 + P_3 \frac{x_1^6}{6} + P_4 x_1 x_2 + P_5 x_2^2$	
	$P_i = 1 \pm 0.5$ $P_{i+d} = 0 \pm 0.4$
$f(x) = \sum_{i=1}^{d} P_i (x_i + P_{i+d})^2$	

<table>
<tr>
<td> 

$f(x) = \sin(x_1 + x_2) + (x_1 - x_2)^2 - P_1 x_1 + P_2 x_2 + P_3$</td>
<td>$P_1 = 1.5 \pm 1$
$P_2 = 2.5 \pm 1$
$P_3 = 1 \pm 0.5$</td>
</tr>
<tr><td></td><td></td></tr>
<tr>
<td> 

$f(x) = -\sum_{i=1}^{4} \alpha_i exp\left(-\sum_{j=1}^{3} P_1 A_{ij} (P_2 x_j - P_3 B_{ij})^2\right)$</td>
<td>$\alpha = (1.0, 1.2, 3.0, 3.2)^T$
$A = \begin{pmatrix} 3.0 & 10 & 30 \\ 0.1 & 10 & 35 \\ 3.0 & 10 & 30 \\ 0.1 & 10 & 35 \end{pmatrix}$
$B = 10^{-4} \begin{pmatrix} 3689 & 1170 & 2673 \\ 4699 & 4387 & 7470 \\ 1091 & 8732 & 5547 \\ 381 & 5743 & 8828 \end{pmatrix}$
$P_1 = 1 \pm 0.6$
$P_2 = 1 \pm 0.2$
$P_3 = 1 \pm 0.2$</td>
</tr>
<tr><td></td><td></td></tr>
</table>

BO, ChIDDO, and MISO algorithm descriptions

Algorithm 1: Bayesian Optimization with no physics model data
Procedure: BO
Input: A set of N_{init} experimental points, X^{exp}, where $X^{exp} \in_R$ DS evaluated to give the objective function value, y^{exp}. $P^{exp} = \{X^{exp}, y^{exp}\}$

1. P^{exp} is passed to the GPR to make predictions, μ^{GPR}, and uncertainties, σ^{GPR}, at any point in DS.
2. P^{exp}, μ^{GPR}, and σ^{GPR} are passed to the acquisition function to select n_b new experiments, X^{next}, to evaluate.
3. Objective function is evaluated at X^{next} to give y^{next} (P^{next}) and then $P^{exp} = P^{exp} \cup P^{next}$.
4. Steps 1-3 is repeated until 50, N_{total}, experiments are evaluated.
5. **Return**: P^{exp}

Algorithm 2: Chemically-informed data driven optimization
Procedure: ChIDDO
Input: A set of N_{init} experimental points, X^{exp}, where $X^{exp} \in_R$ DS evaluated to give the objective function value, y^{exp}. $P^{exp} = \{X^{exp}, y^{exp}\}$
Input: A model, M, of the system under study to approximate the objective function

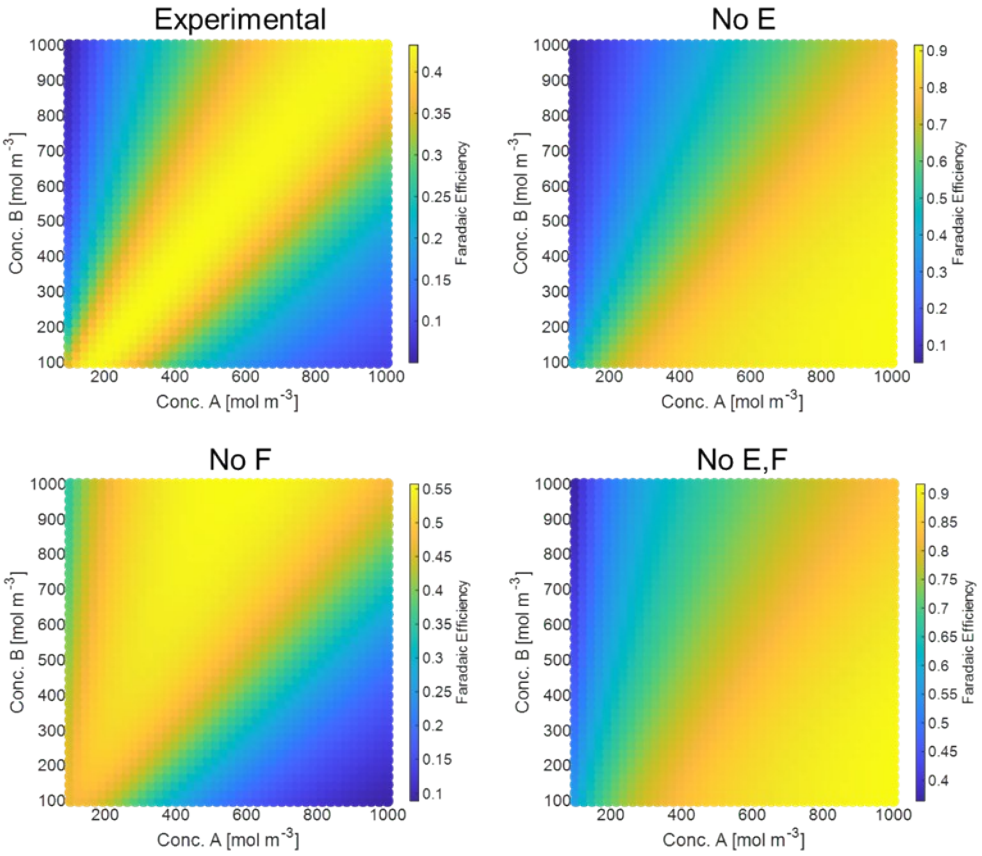
1. $X^{phys} \in_R$ DS are evaluated by M to give y_{phys} ($P^{phys} = \{X^{phys}, y^{phys}\}$) and then $P^{tot} = P^{exp} \cup P^{phys}$
2. P^{tot} is passed to the GPR to make predictions, μ^{GPR}, and uncertainties, σ^{GPR}, at any point in DS.
3. P^{exp}, μ^{GPR}, and σ^{GPR} are passed to the acquisition function to select n_b new experiments, X^{next}, to evaluate.
4. Objective function is evaluated at X^{next} to give y^{next} (P^{next}) and then $P^{exp} = P^{exp} \cup P^{next}$.
5. P^{exp} is used to refine the parameters of M by using non-linear regression.
6. Steps 1-5 is repeated until 50, N_{total}, experiments are evaluated.
7. **Return**: P^{exp}

Algorithm 3: Multi-information Source Optimization as reported in Poloczek et al.[2]
Procedure: misoKG
Input: A set of N_{init} experimental points, X^{exp}, where $X^{exp} \in_R$ DS evaluated to give the objective function value, y^{exp}. $P^{exp} = \{X^{exp}, y^{exp}\}$
Input: A set of n_l information sources, l, that give different biases on the objective function.

1. $X^{exp} \in_R$ DS are evaluated by each l to give $y^{exp}(l, X^{exp})$.
2. X^{exp} and y^{exp} are passed to the GPR to make the posterior prediction.
3. Until the budget for samples is consumed, determine the l and X^{next} that maximize the misoKG factor (Equation 1 in Reference 1)[2].
4. Objective function is evaluated at X^{next} and l and the posterior is updated with the new information.
5. Steps 2-4 are repeated until sample budget is exhausted.

Return: P^{exp}

Comparison between 2D simplified models and experimental objective function



BO and ChIDDO comparison of d_x to Edisonian search using MRB

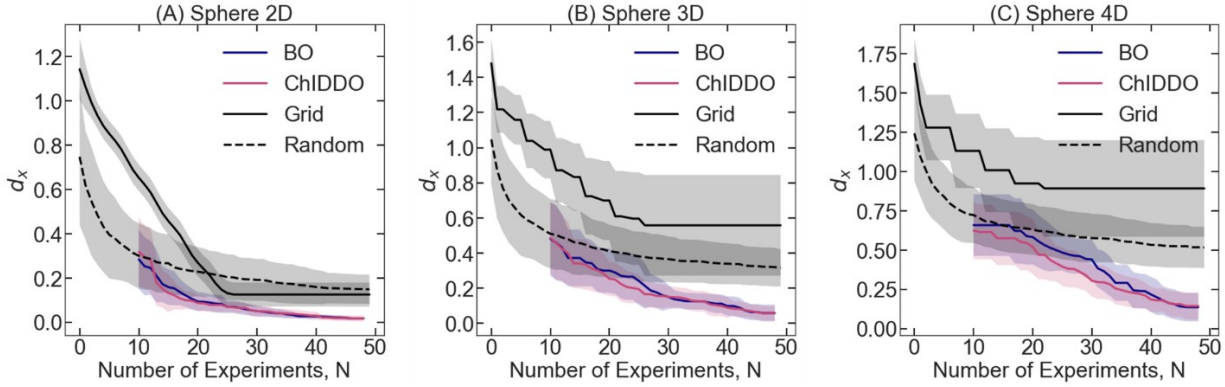


Figure S1. d_x versus number of experiments, N, comparing BO and ChIDDO with the Edisonian random and grid search. (A) 2D Sphere, (B) 3D Sphere, (C) 4D Sphere. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. For each of the BO/ChIDDO experiments, the MRB acquisition function was used.

BO and ChIDDO comparison to Edisonian search using MRB

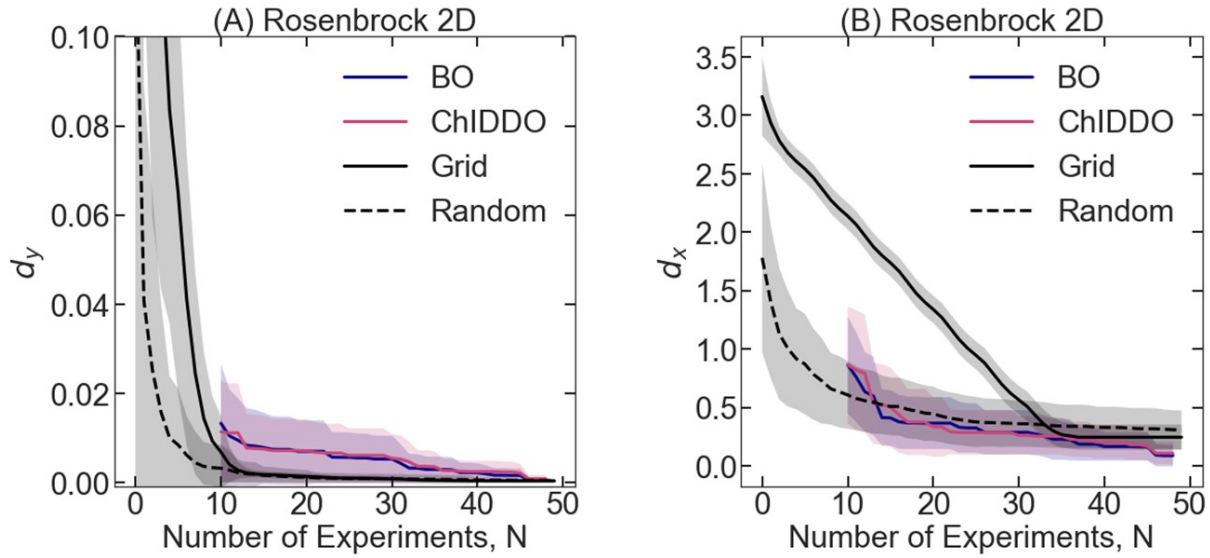


Figure S2. d_y and d_x versus number of experiments, N, comparing BO and ChIDDO with the Edisonian random and grid search on the 2D Rosenbrock function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. For each of the BO/ChIDDO experiments, the MRB acquisition function was used.

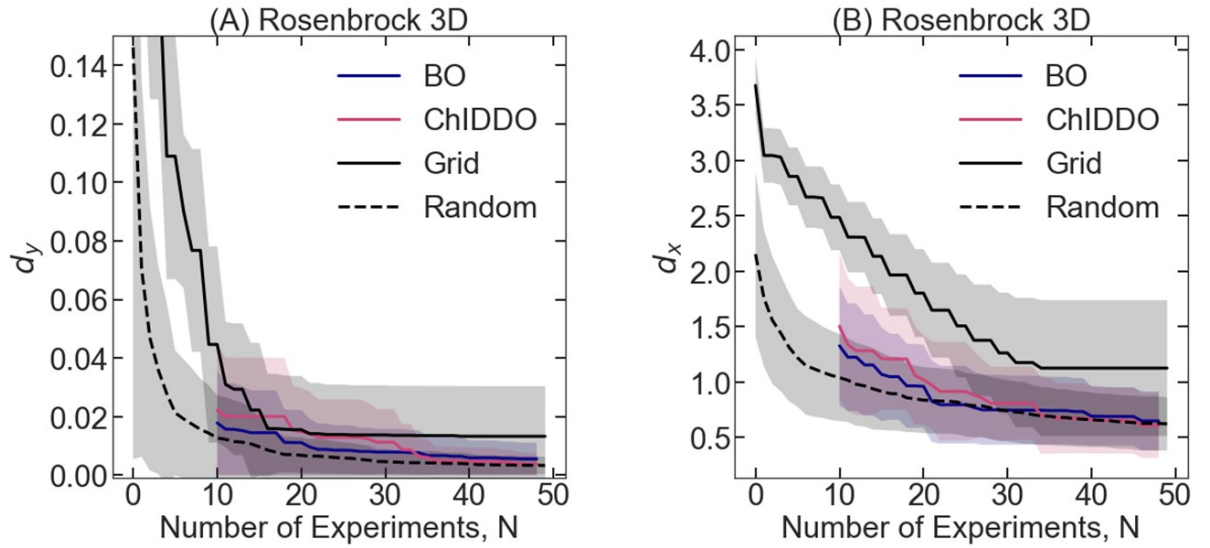


Figure S3. d_y and d_x versus number of experiments, N, comparing BO and ChIDDO with the Edisonian random and grid search on the 3D Rosenbrock function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. For each of the BO/ChIDDO experiments, the MRB acquisition function was used.

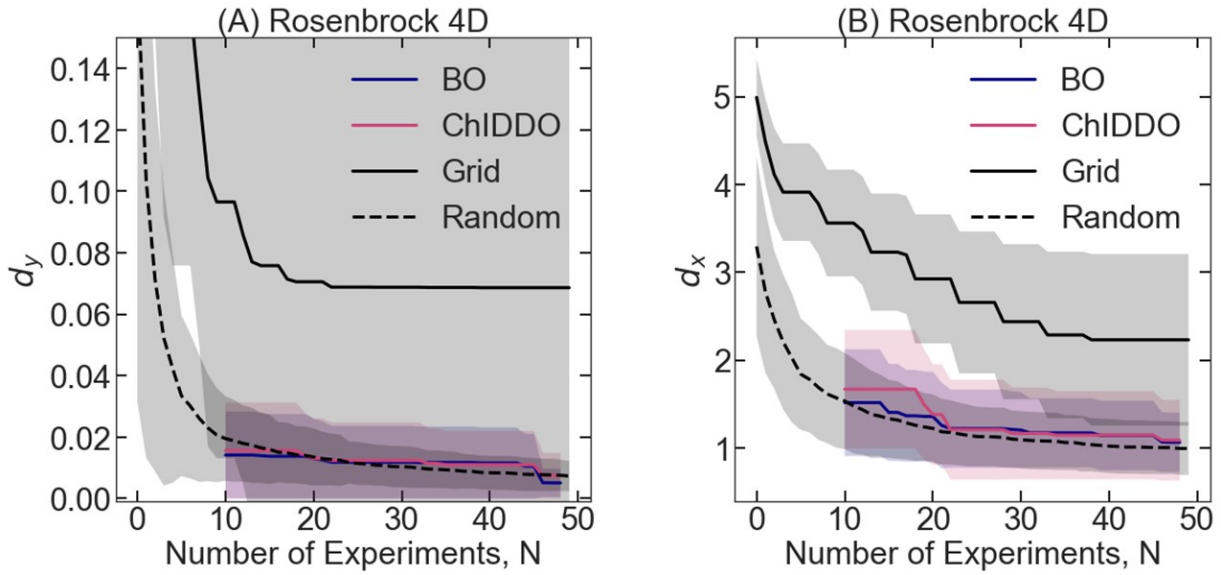


Figure S4. d_y and d_x versus number of experiments, N, comparing BO and ChIDDO with the Edisonian random and grid search on the 4D Rosenbrock function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. For each of the BO/ChIDDO experiments, the MRB acquisition function was used.

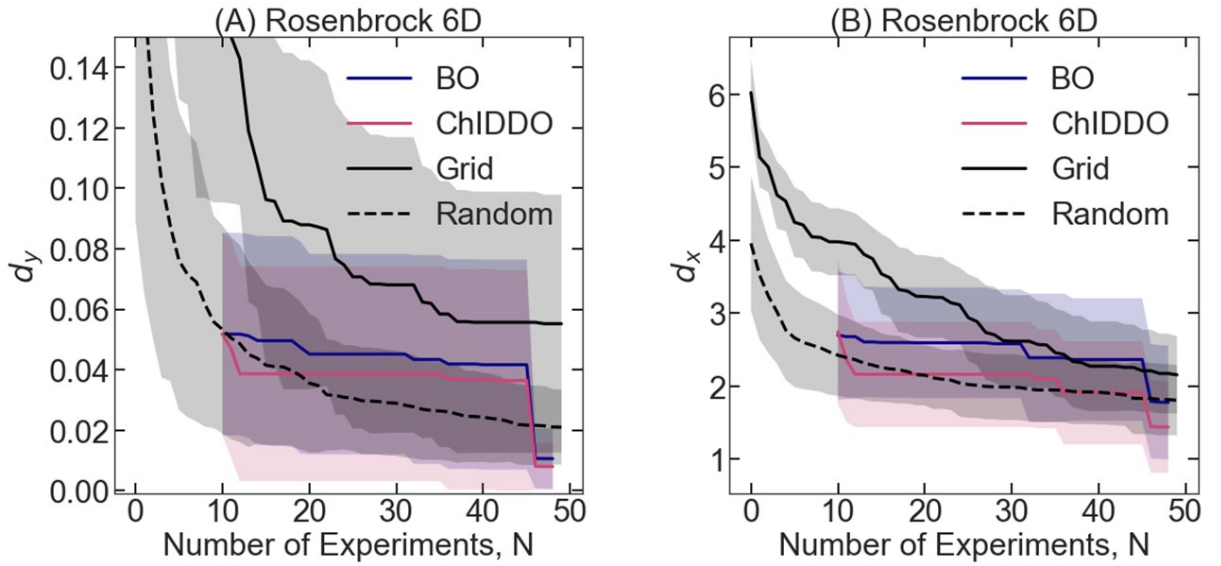


Figure S5. d_y and d_x versus number of experiments, N, comparing BO and ChIDDO with the Edisonian random and grid search on the 6D Rosenbrock function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. For each of the BO/ChIDDO experiments, the MRB acquisition function was used.

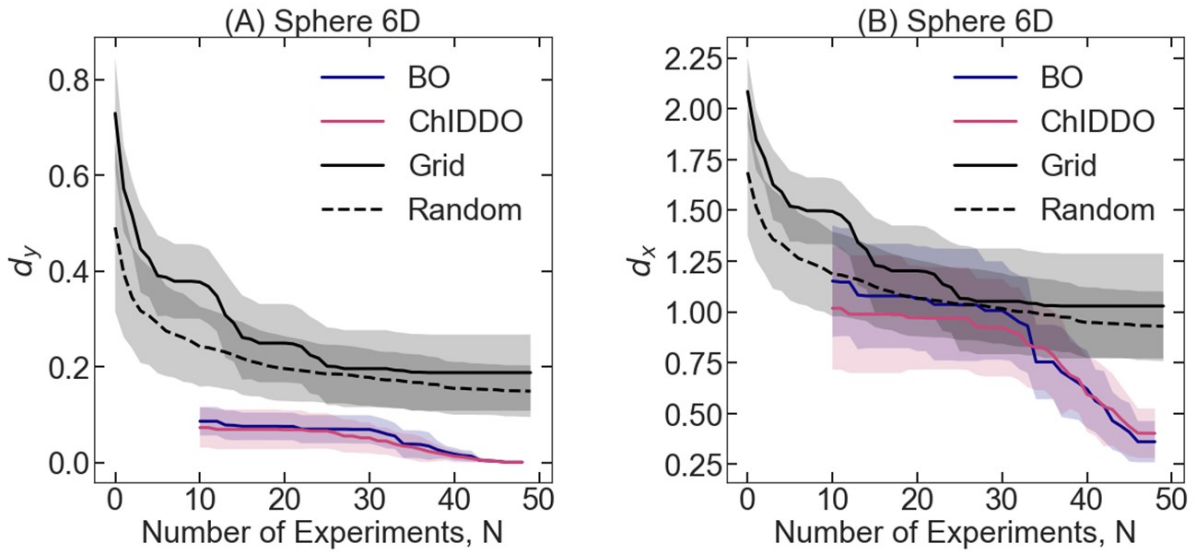


Figure S6. d_y and d_x versus number of experiments, N, comparing BO and ChIDDO with the Edisonian random and grid search on the 6D Sphere function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. For each of the BO/ChIDDO experiments, the MRB acquisition function was used.

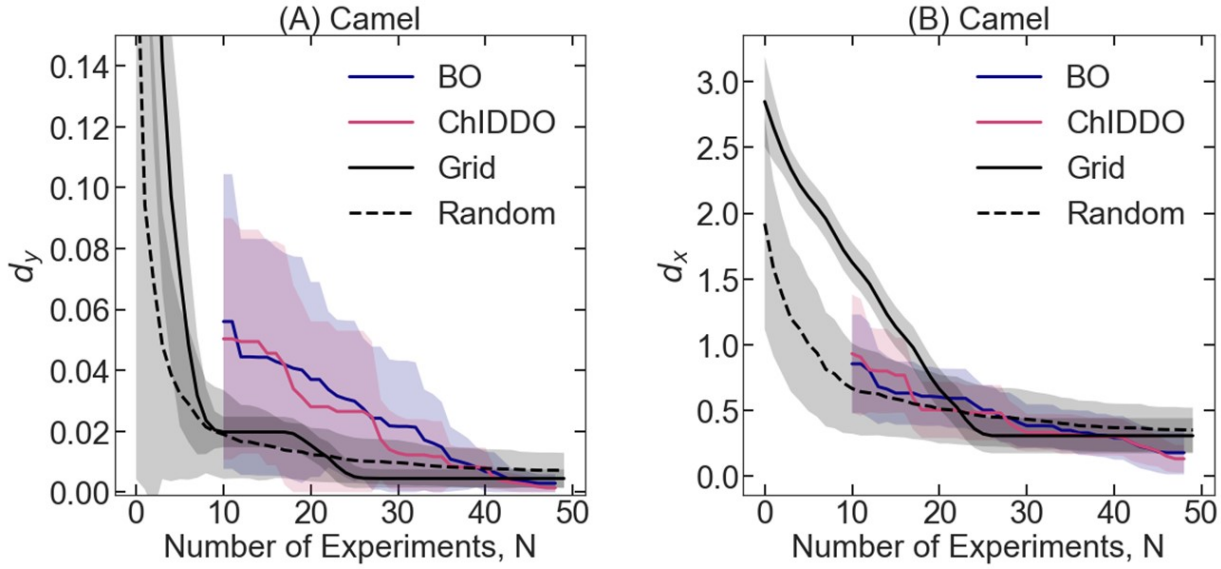


Figure S7. d_y and d_x versus number of experiments, N, comparing BO and ChIDDO with the Edisonian random and grid search on the 2D Camel function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. For each of the BO/ChIDDO experiments, the MRB acquisition function was used.

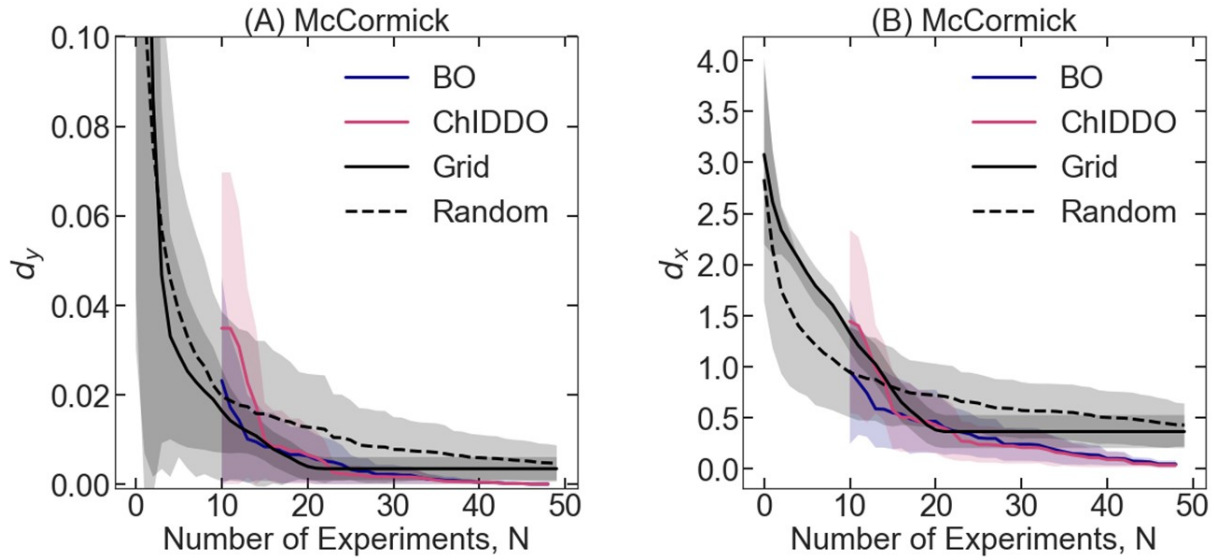


Figure S8. d_y and d_x versus number of experiments, N, comparing BO and ChIDDO with the Edisonian random and grid search on the 2D Mccormick function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. For each of the BO/ChIDDO experiments, the MRB acquisition function was used.

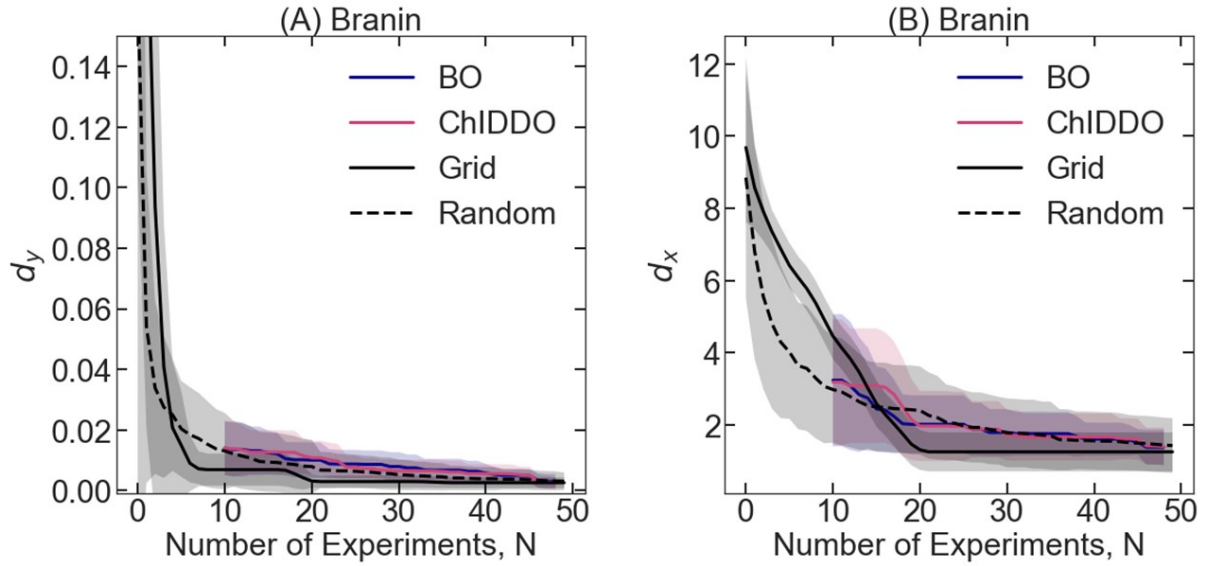


Figure S9. d_y and d_x versus number of experiments, N, comparing BO and ChIDDO with the Edisonian random and grid search on the 2D Branin function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. For each of the BO/ChIDDO experiments, the MRB acquisition function was used.

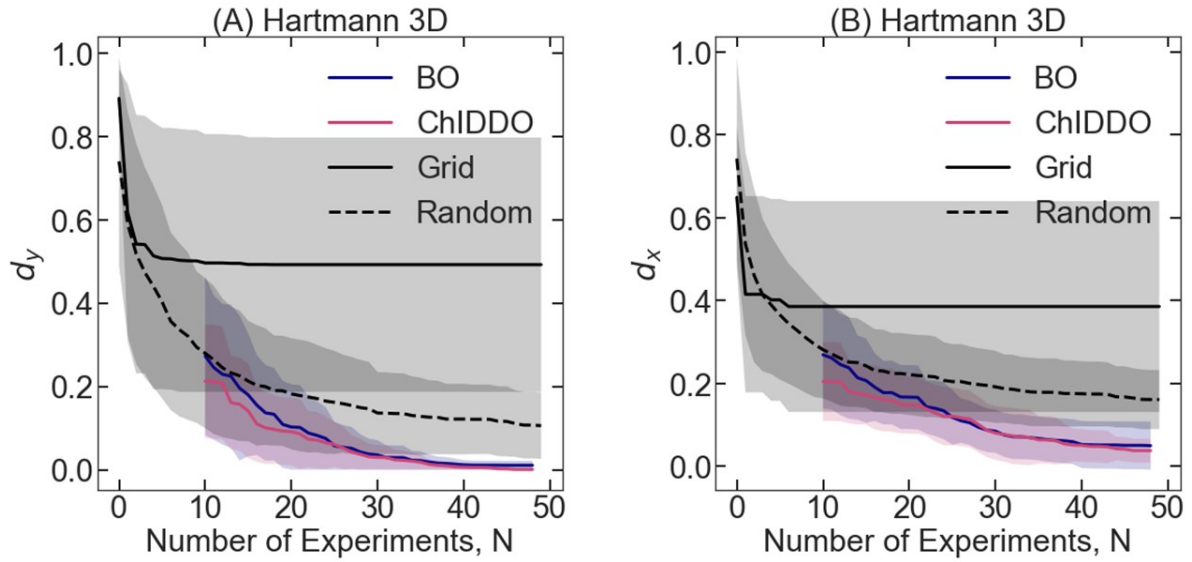


Figure S10. d_y and d_x versus number of experiments, N, comparing BO and ChIDDO with the Edisonian random and grid search on the 3D Hartmann function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. For each of the BO/ChIDDO experiments, the MRB acquisition function was used.

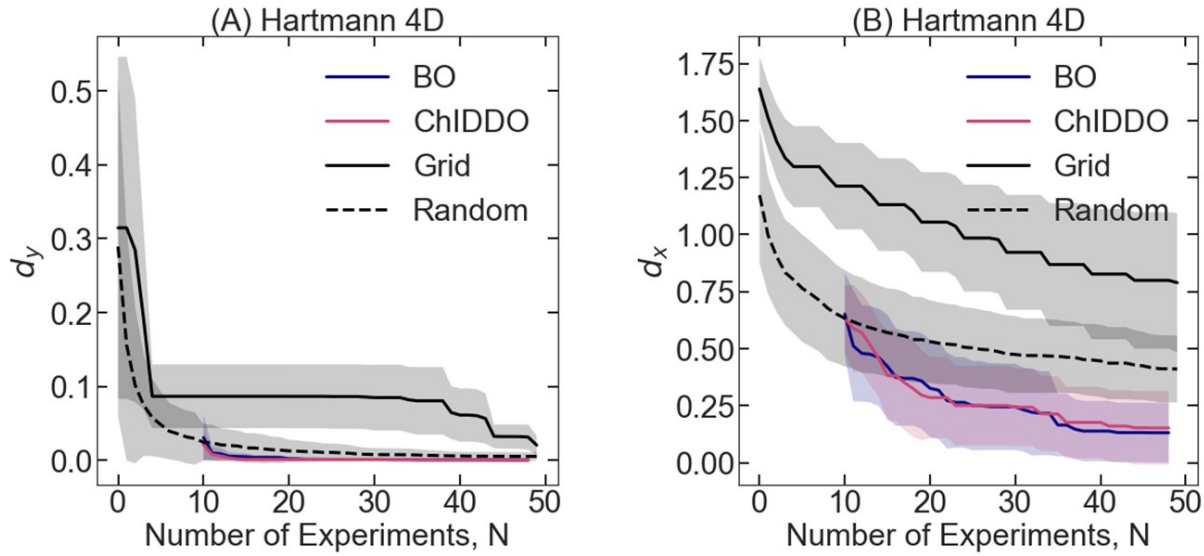


Figure S11. d_y and d_x versus number of experiments, N, comparing BO and ChIDDO with the Edisonian random and grid search on the 4D Hartmann function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. For each of the BO/ChIDDO experiments, the MRB acquisition function was used.

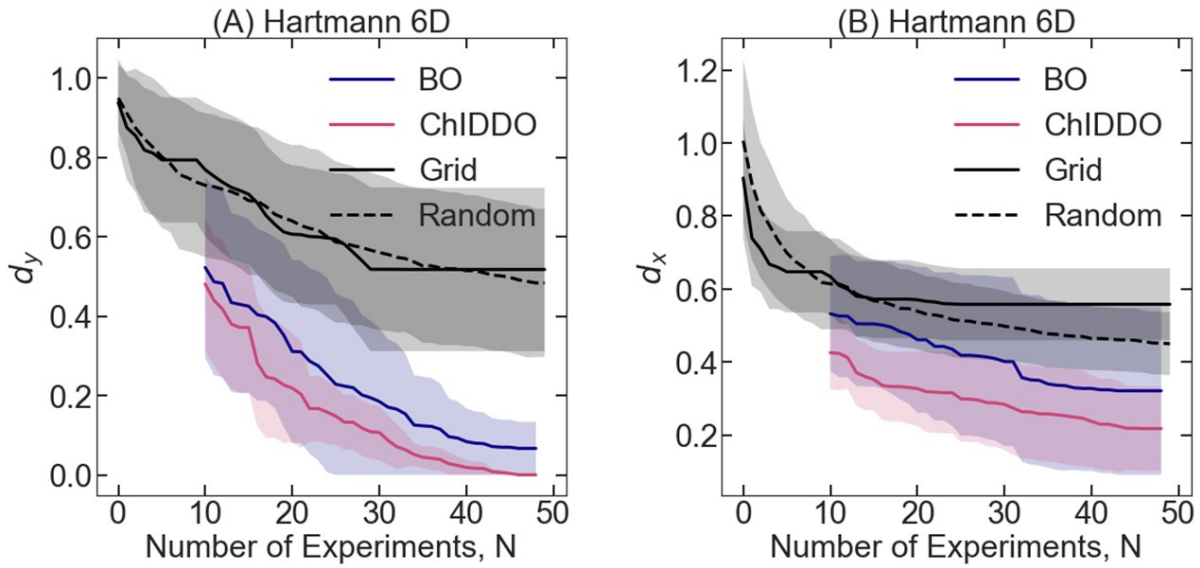


Figure S12. d_y and d_x versus number of experiments, N, comparing BO and ChIDDO with the Edisonian random and grid search on the 6D Hartmann function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. For each of the BO/ChIDDO experiments, the MRB acquisition function was used.

Acquisition function comparison of d_x for Hartmann objective functions

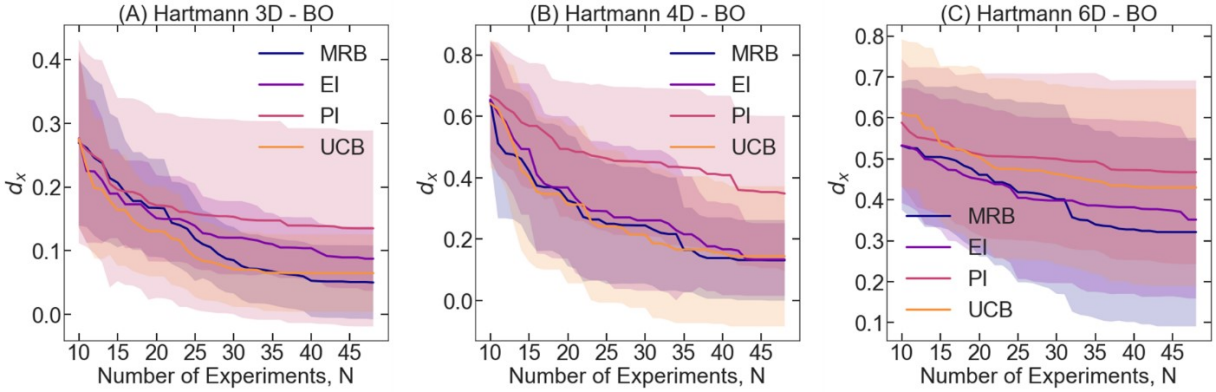


Figure S13. d_x versus number of experiments, N, comparing different acquisition functions using BO on: (A) 3D Hartmann, (B) 4D Hartmann, (C) 6D Hartmann. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

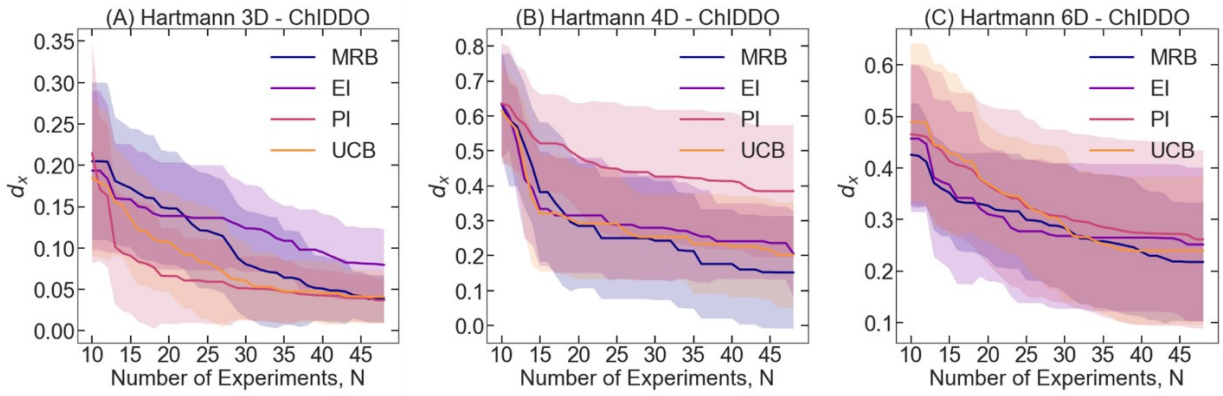


Figure S14. d_x versus number of experiments, N, comparing different acquisition functions using ChIDDO on: (A) 3D Hartmann, (B) 4D Hartmann, (C) 6D Hartmann. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

Acquisition function comparison for all objective functions

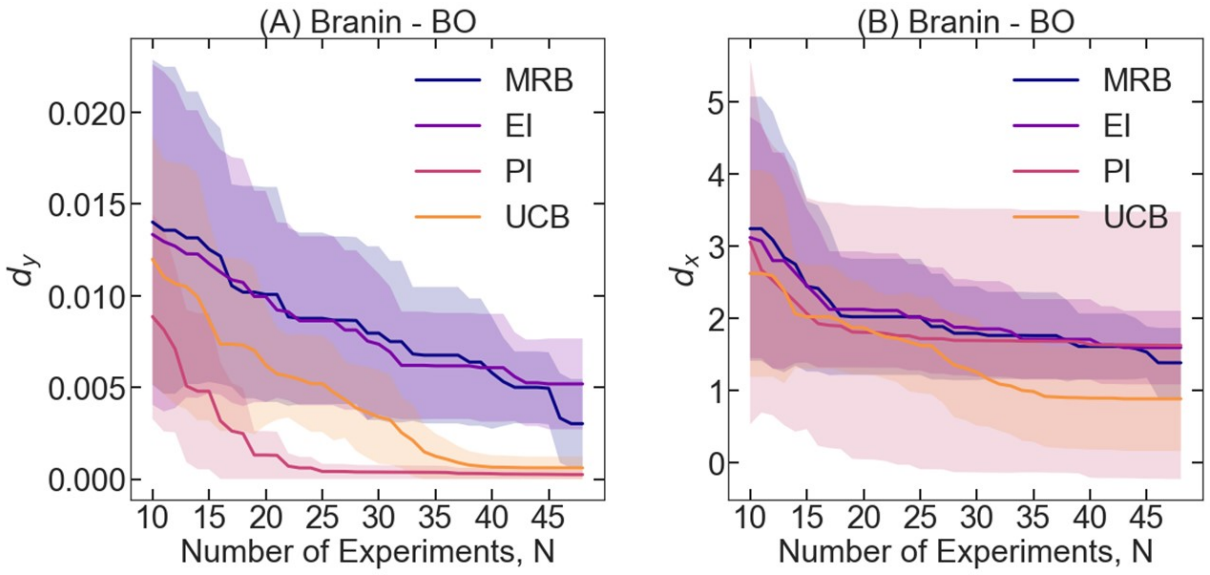


Figure S15. d_y and d_x versus number of experiments, N, comparing different acquisition functions using BO on the 2D Branin function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

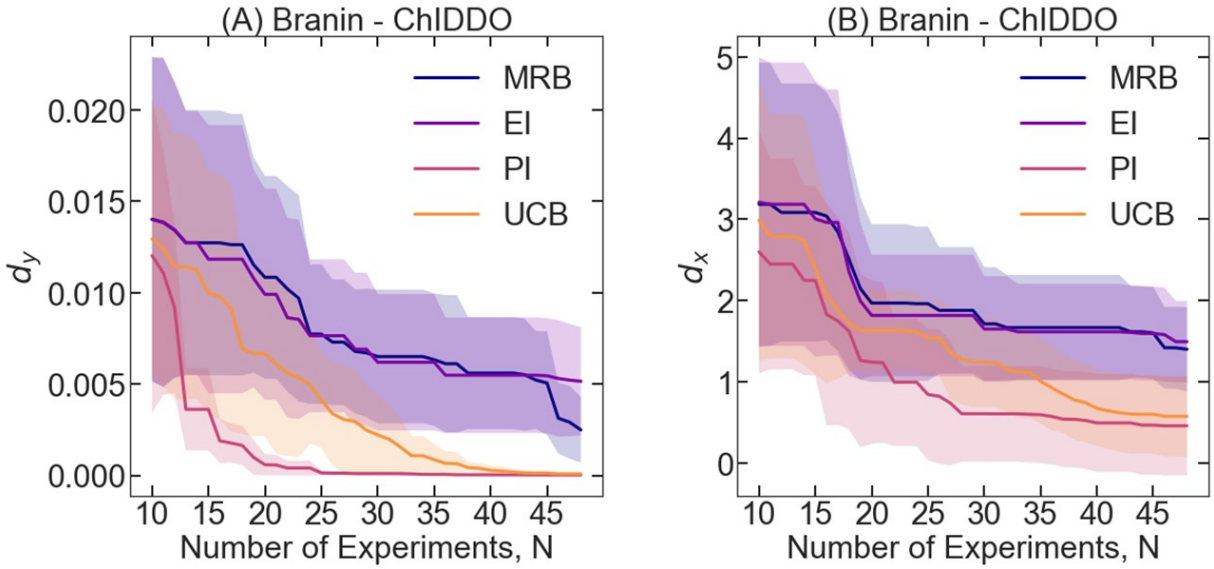


Figure S16. d_y and d_x versus number of experiments, N, comparing different acquisition functions using ChIDDO on the 2D Branin function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

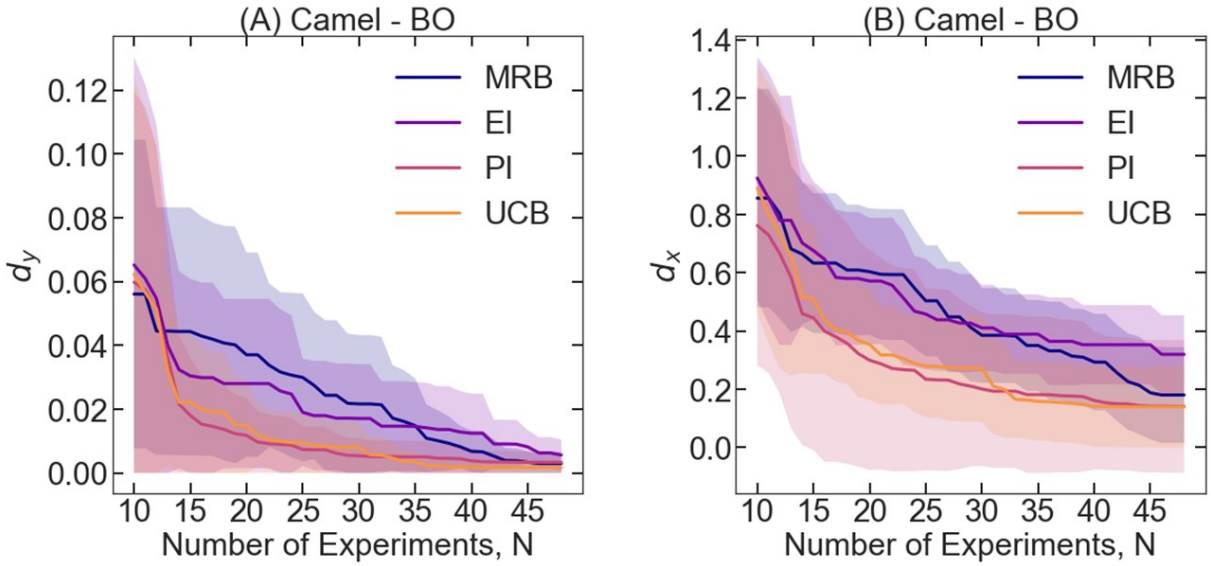


Figure S17. d_y and d_x versus number of experiments, N, comparing different acquisition functions using BO on the 2D Camel function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

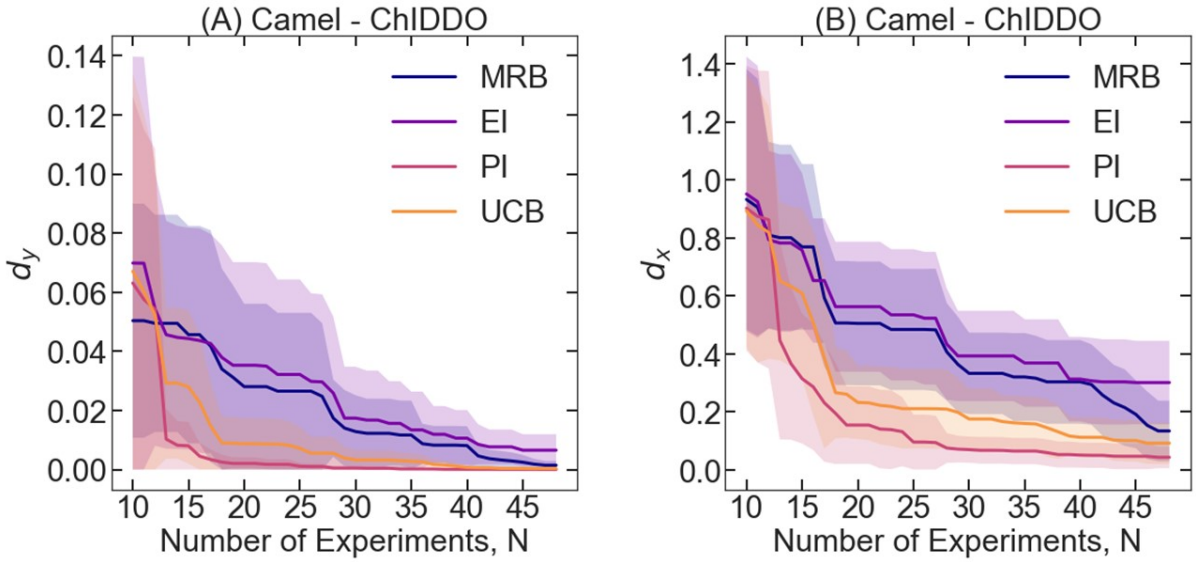


Figure S18. d_y and d_x versus number of experiments, N, comparing different acquisition functions using ChIDDO on the 2D Camel function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

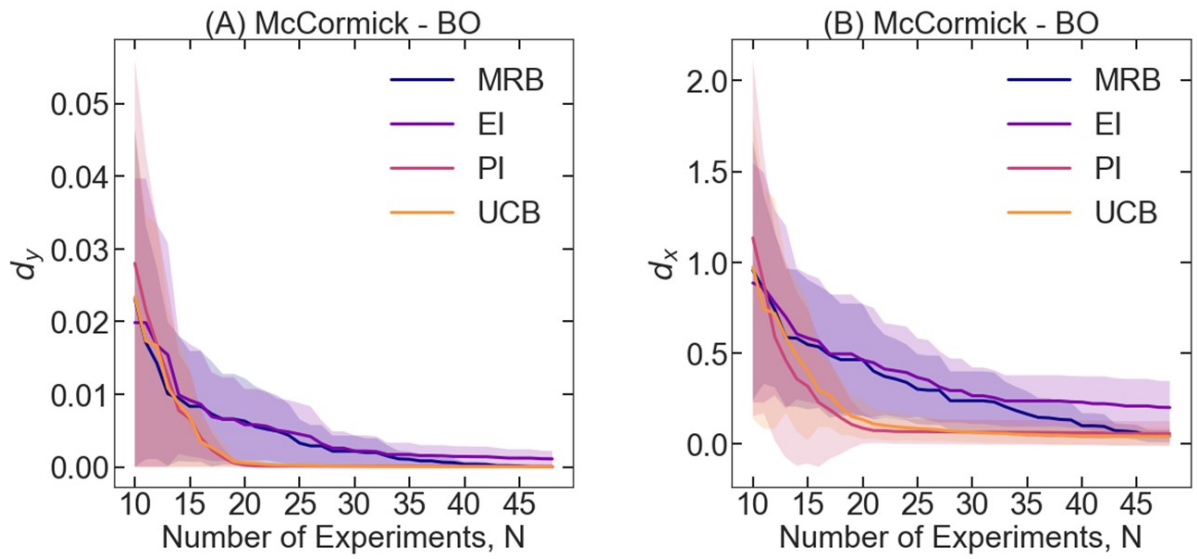


Figure S19. d_y and d_x versus number of experiments, N, comparing different acquisition functions using BO on the 2D Mccormic function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

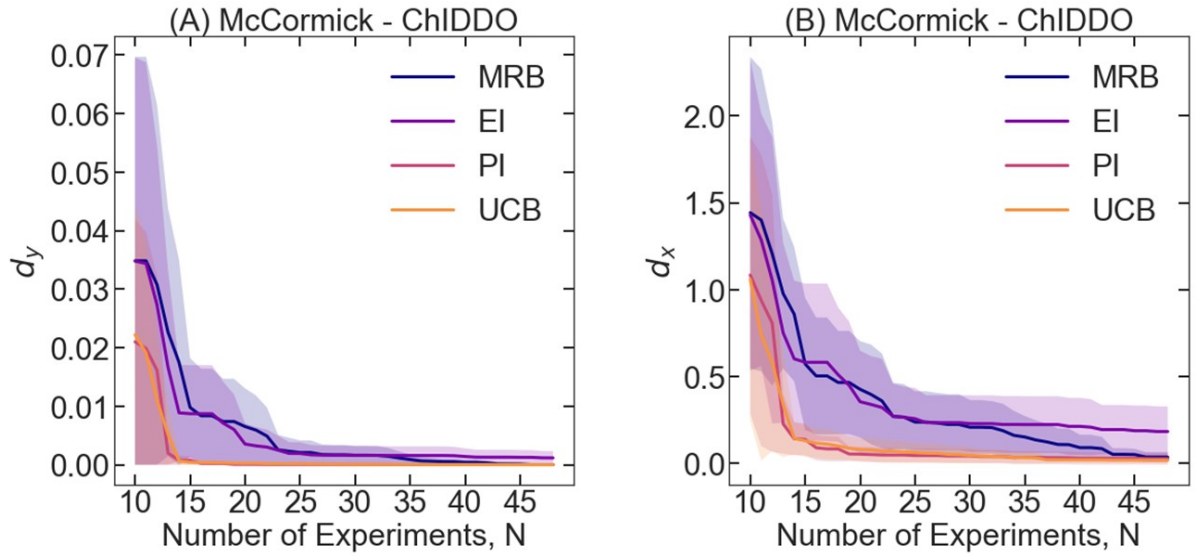


Figure S20. d_y and d_x versus number of experiments, N, comparing different acquisition functions using ChIDDO on the 2D Mccormick function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

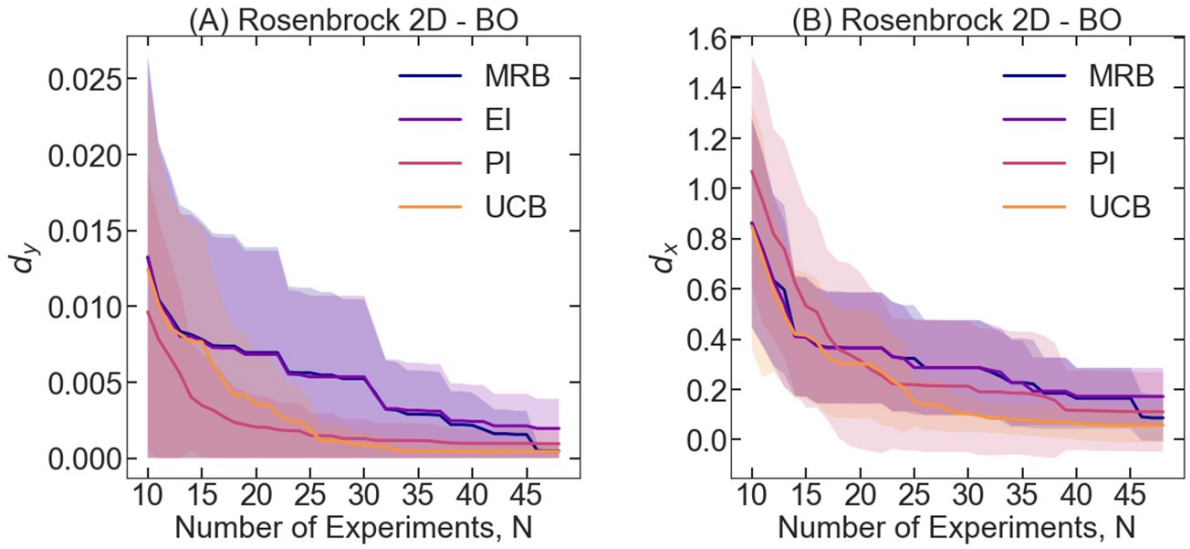


Figure S21. d_y and d_x versus number of experiments, N, comparing different acquisition functions using BO on the 2D Rosenbrock function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

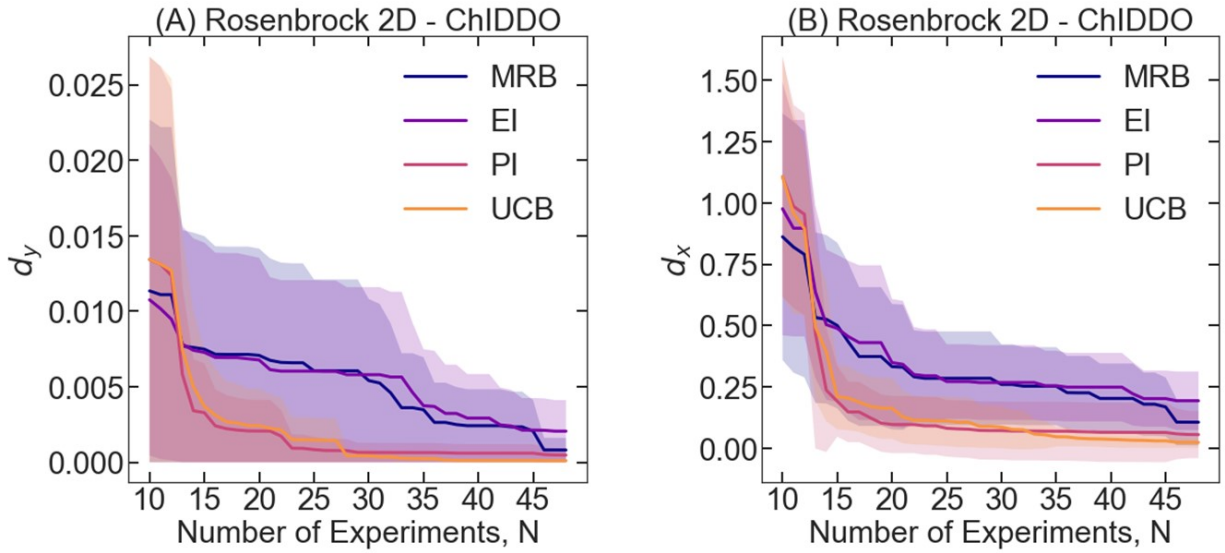


Figure S22. d_y and d_x versus number of experiments, N, comparing different acquisition functions using ChIDDO on the 2D Rosenbrock function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

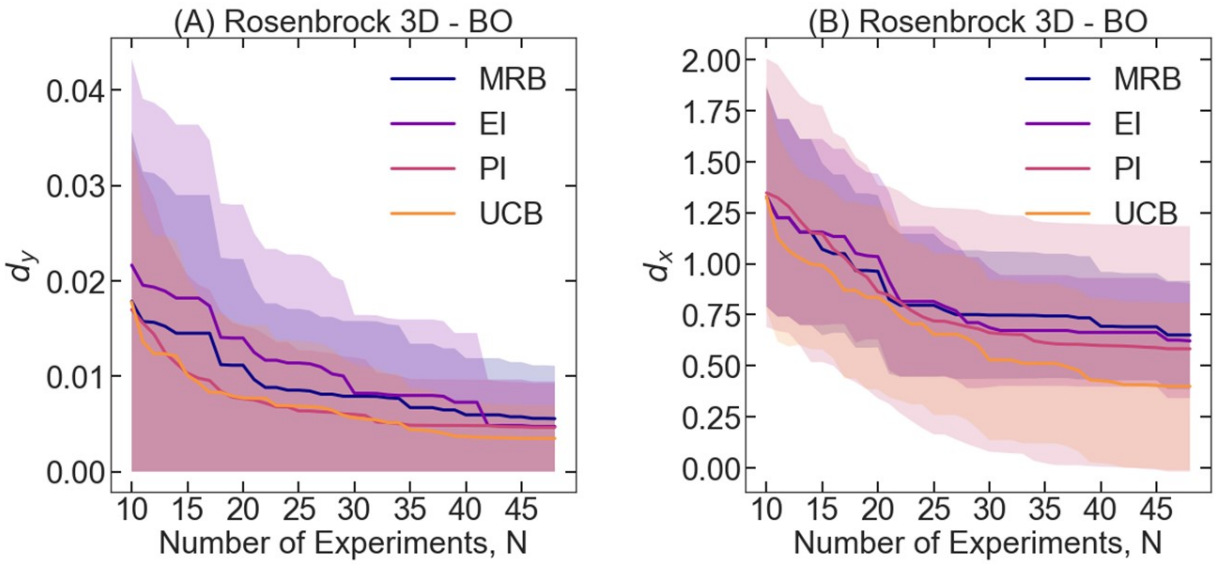


Figure S23. d_y and d_x versus number of experiments, N, comparing different acquisition functions using BO on the 3D Rosenbrock function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

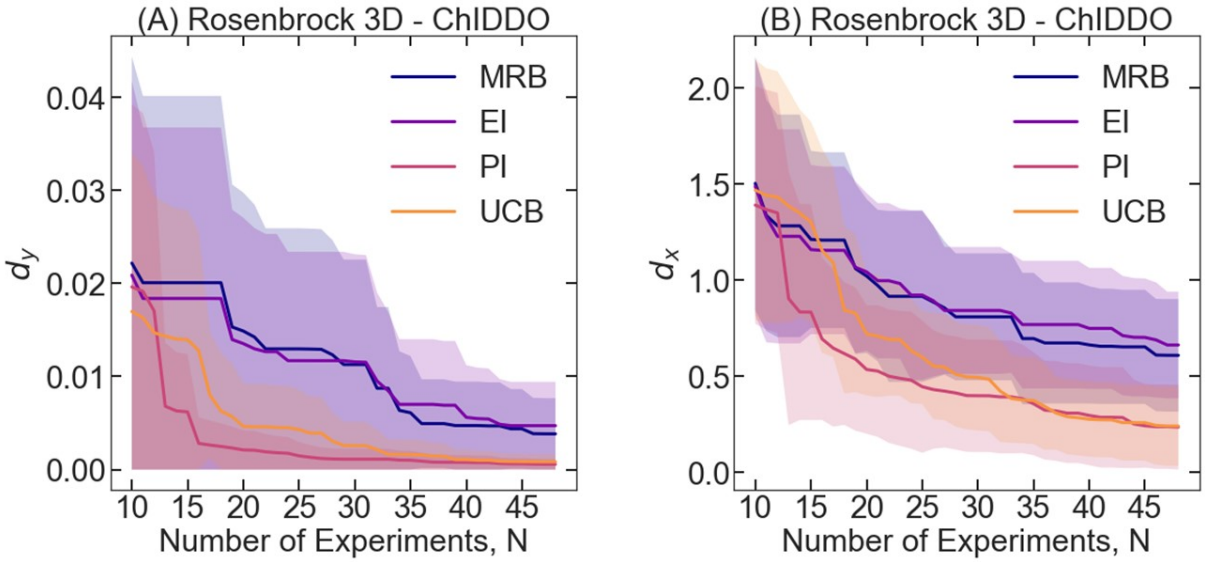


Figure S24. d_y and d_x versus number of experiments, N, comparing different acquisition functions using ChIDDO on the 3D Rosenbrock function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

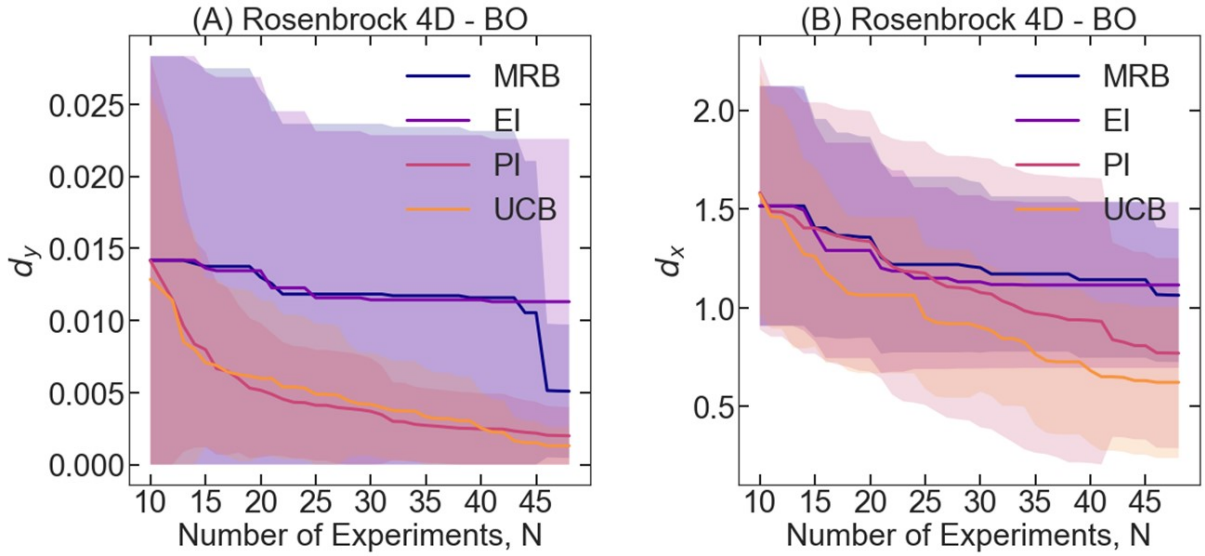


Figure S25. d_y and d_x versus number of experiments, N, comparing different acquisition functions using BO on the 4D Rosenbrock function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

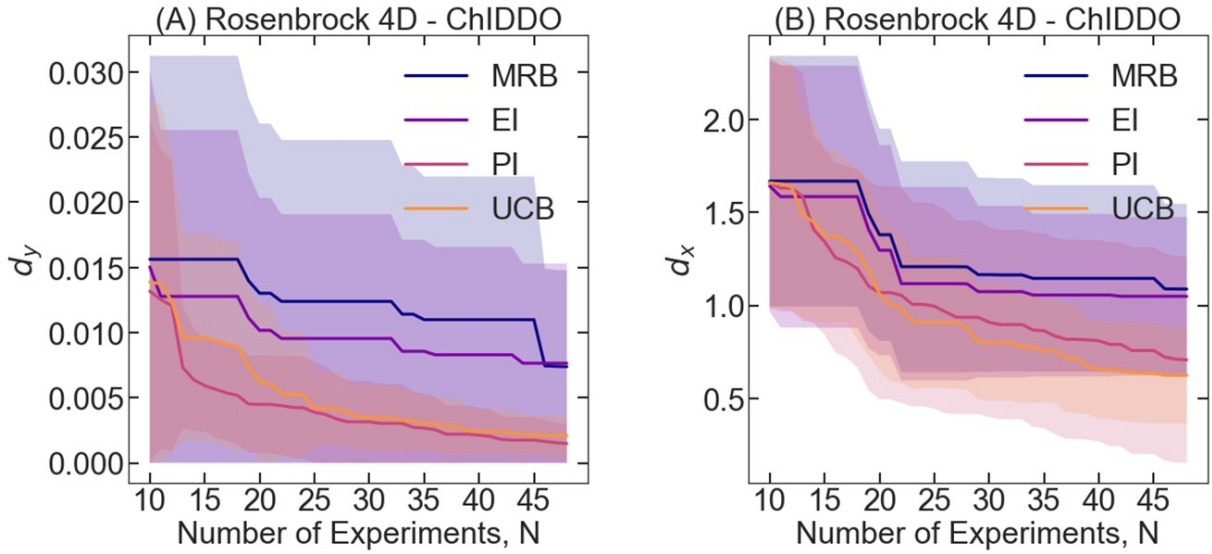


Figure S26. d_y and d_x versus number of experiments, N, comparing different acquisition functions using ChIDDO on the 4D Rosenbrock function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

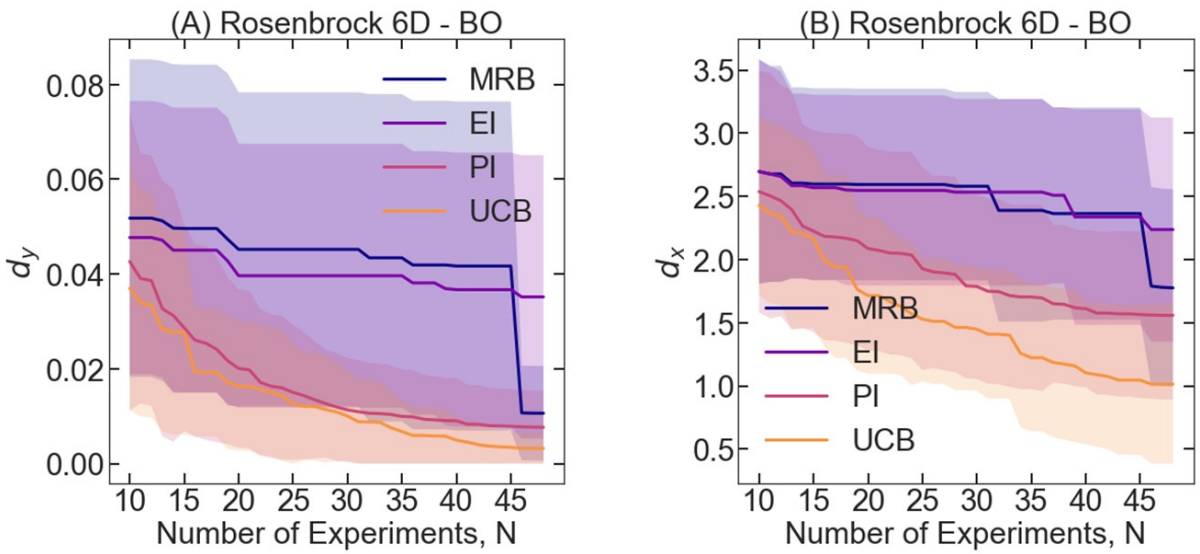


Figure S27. d_y and d_x versus number of experiments, N, comparing different acquisition functions using BO on the 6D Rosenbrock function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

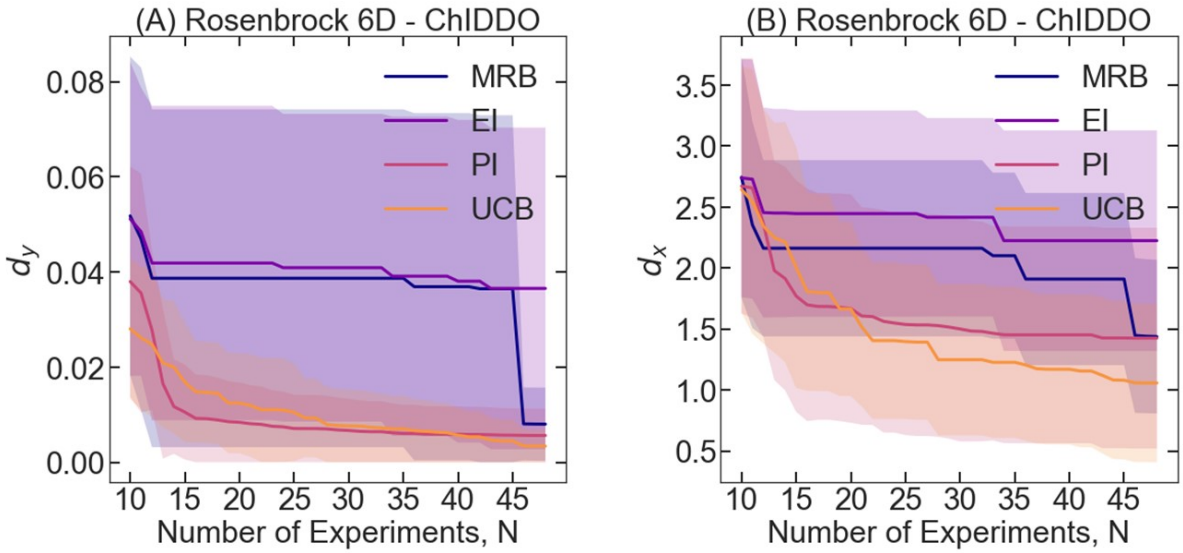


Figure S28. d_y and d_x versus number of experiments, N, comparing different acquisition functions using ChIDDO on the 6D Rosenbrock function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

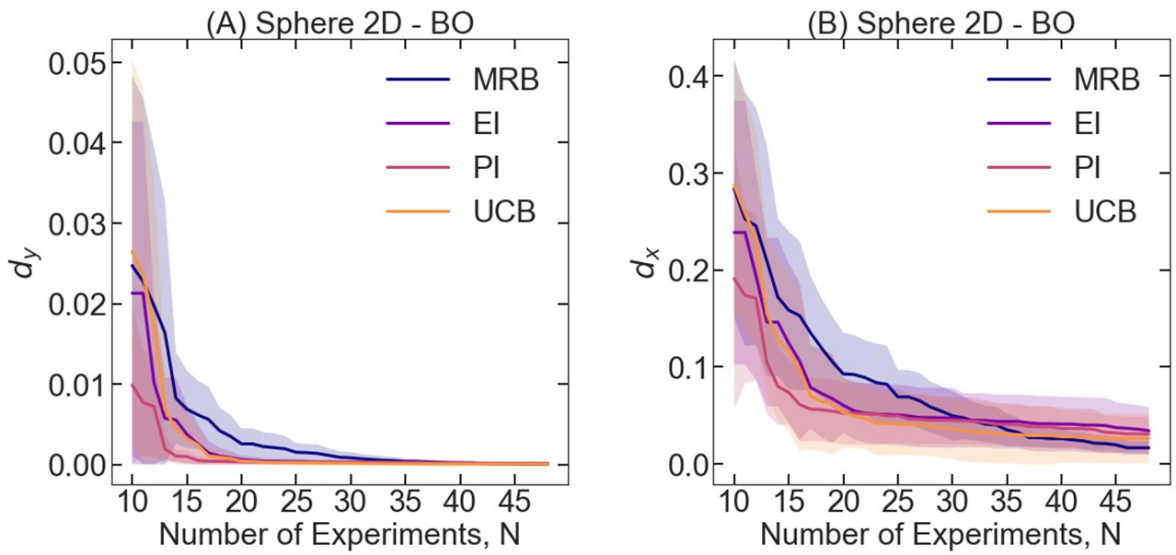


Figure S29. d_y and d_x versus number of experiments, N, comparing different acquisition functions using BO on the 2D Sphere function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

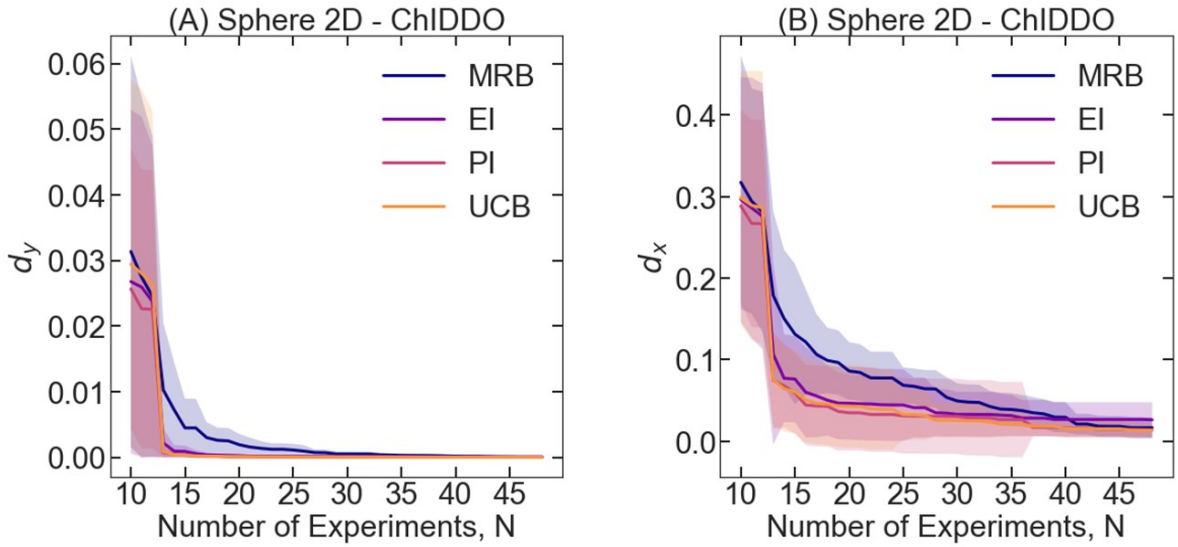


Figure S30. d_y and d_x versus number of experiments, N, comparing different acquisition functions using ChIDDO on the 2D Sphere function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

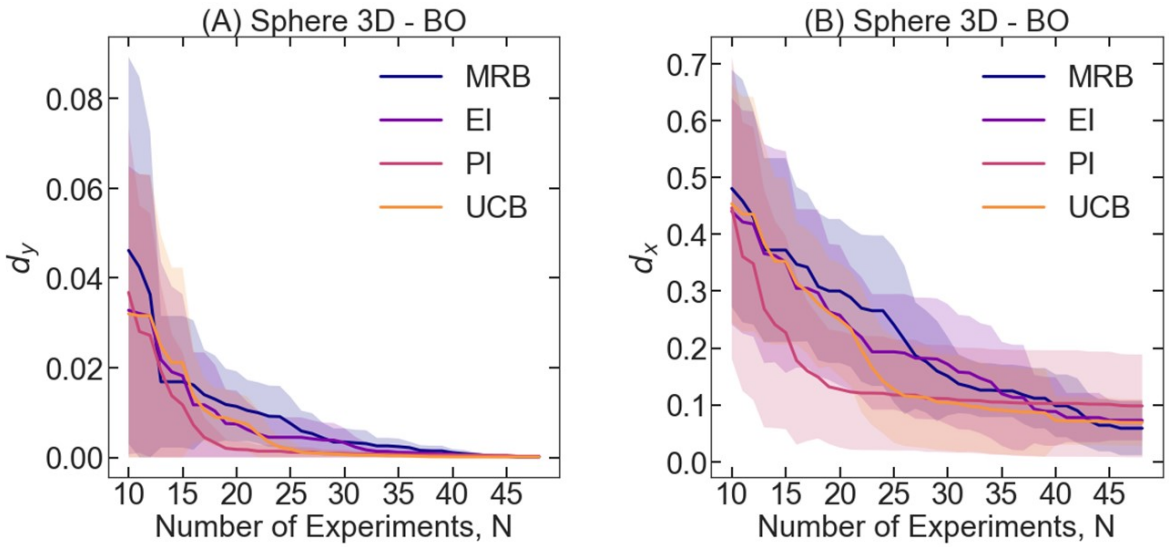


Figure S31. d_y and d_x versus number of experiments, N, comparing different acquisition functions using BO on the 3D Sphere function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

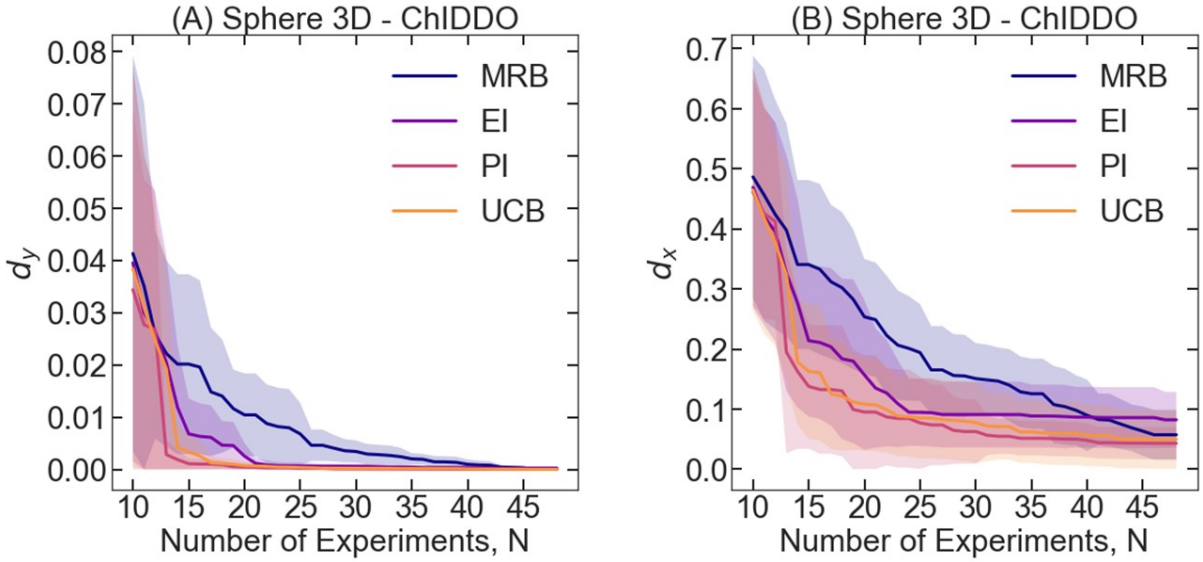


Figure S32. d_y and d_x versus number of experiments, N, comparing different acquisition functions using ChIDDO on the 3D Sphere function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

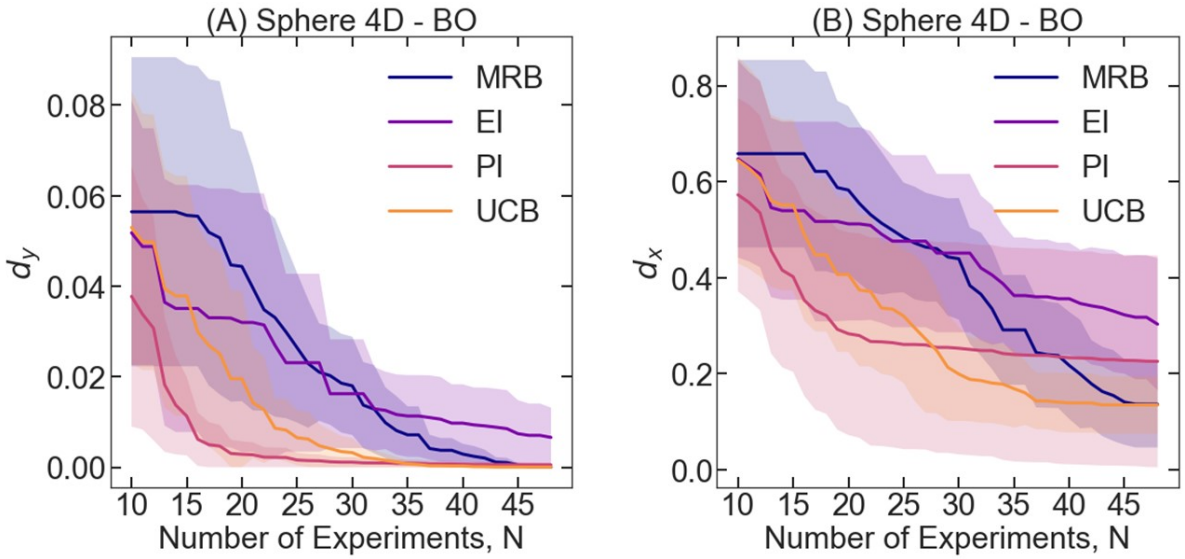


Figure S33. d_y and d_x versus number of experiments, N, comparing different acquisition functions using BO on the 4D Sphere function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

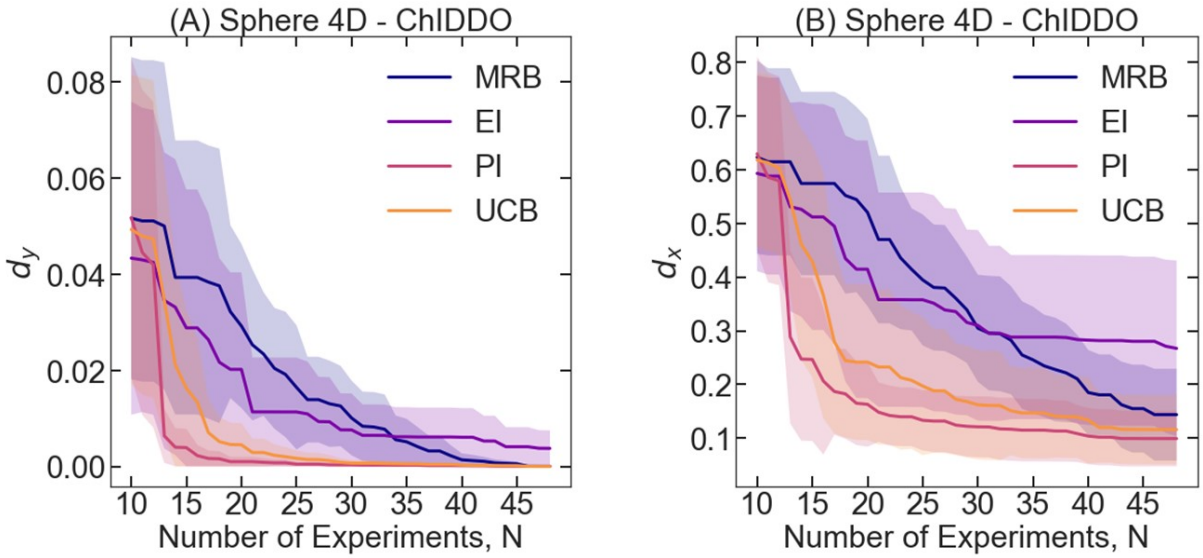


Figure S34. d_y and d_x versus number of experiments, N, comparing different acquisition functions using ChIDDO on the 4D Sphere function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

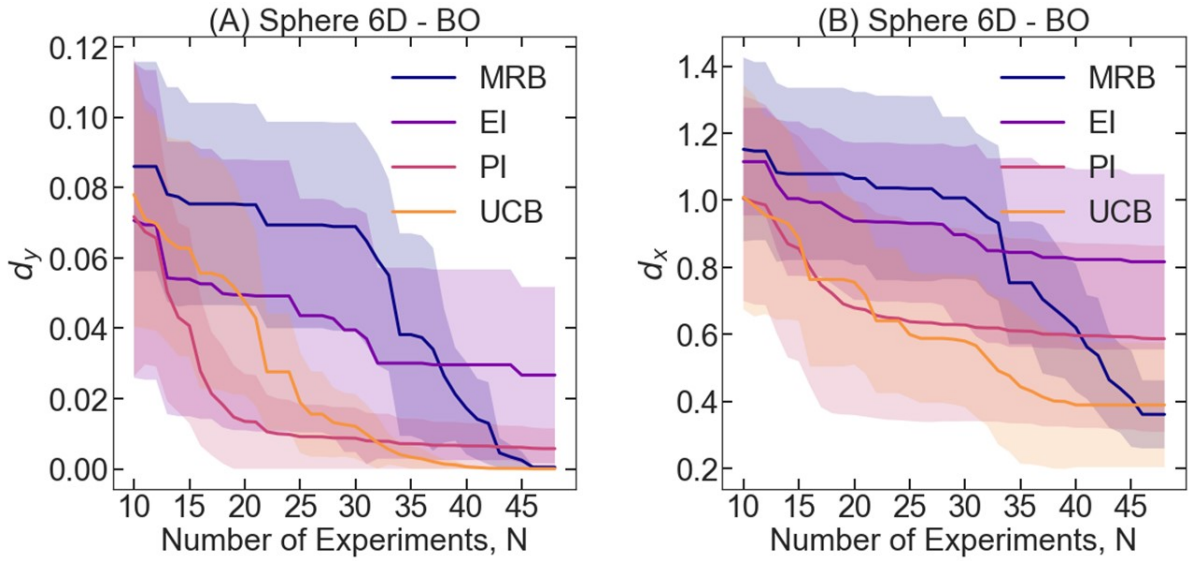


Figure S35. d_y and d_x versus number of experiments, N, comparing different acquisition functions using BO on the 6D Sphere function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

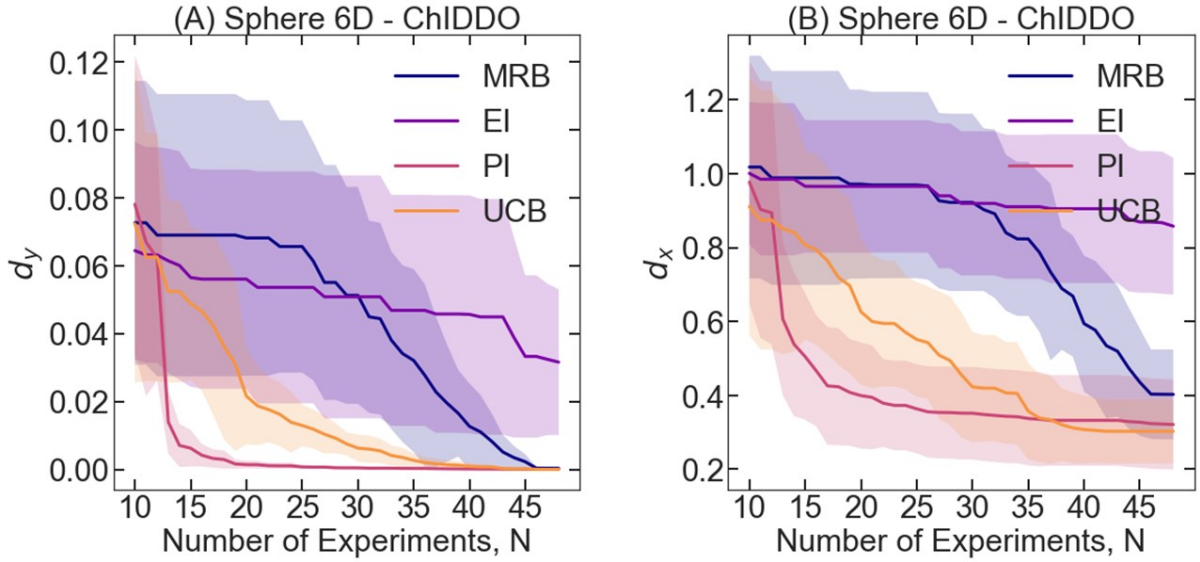


Figure S36. d_y and d_x versus number of experiments, N, comparing different acquisition functions using ChIDDO on the 6D Sphere function. (A) d_y, (B) d_x,. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation.

Noise level comparison for other objective functions using MRB

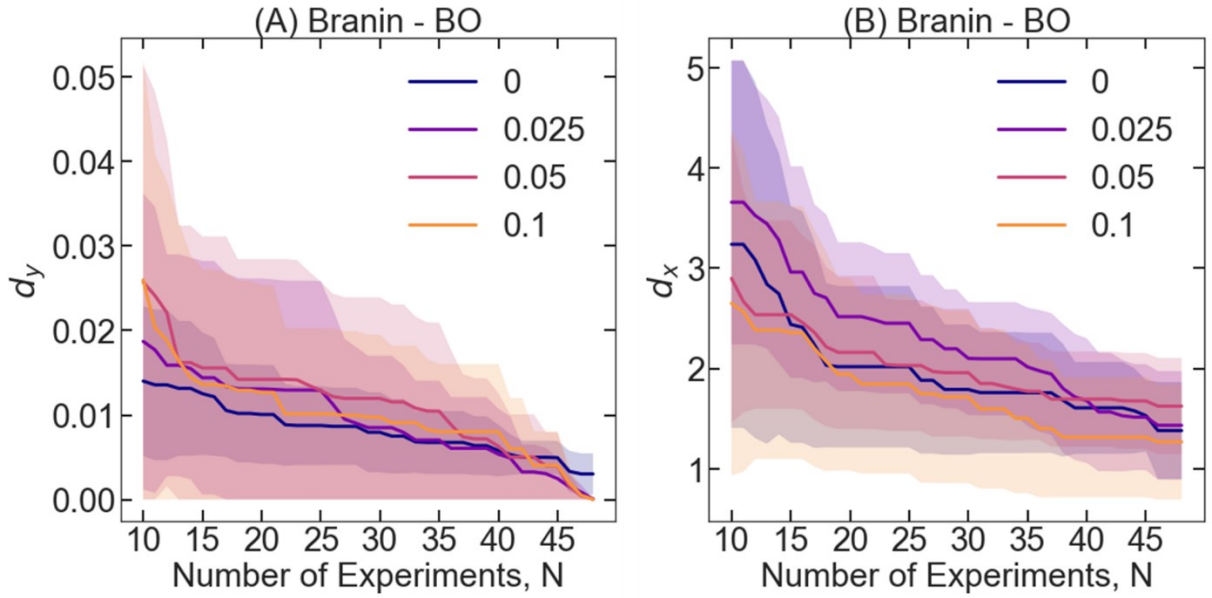


Figure S37. d_y and d_x versus number of experiments, N, comparing different noise levels using BO on the 2D Branin function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

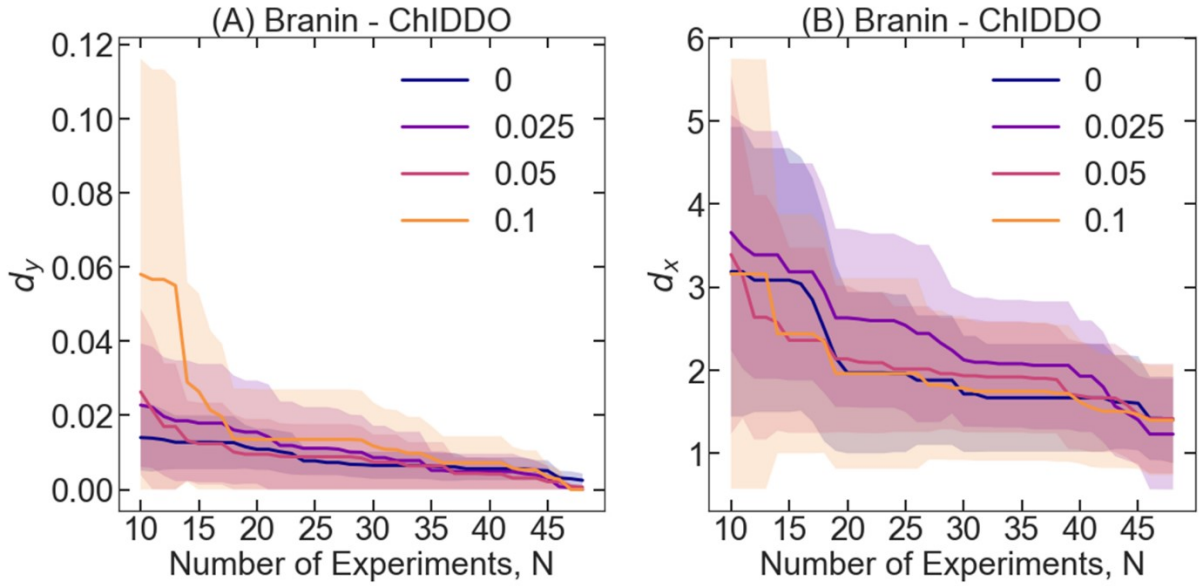


Figure S38. d_y and d_x versus number of experiments, N, comparing different noise levels using ChIDDO on the 2D Branin function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

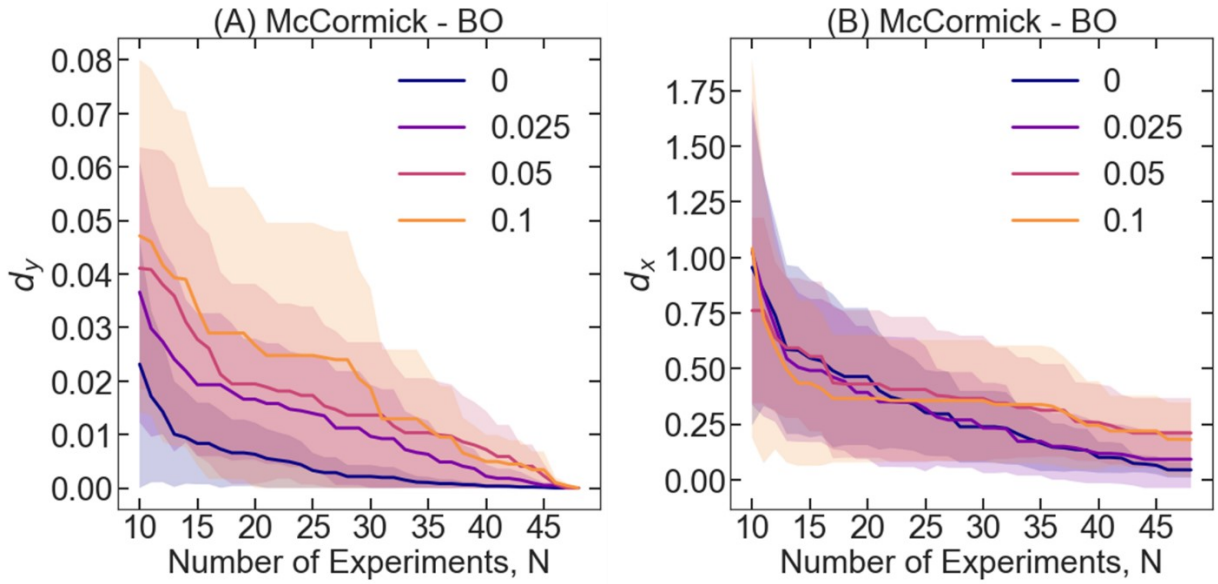


Figure S39. d_y and d_x versus number of experiments, N, comparing different noise levels using BO on the 2D Mccormick function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

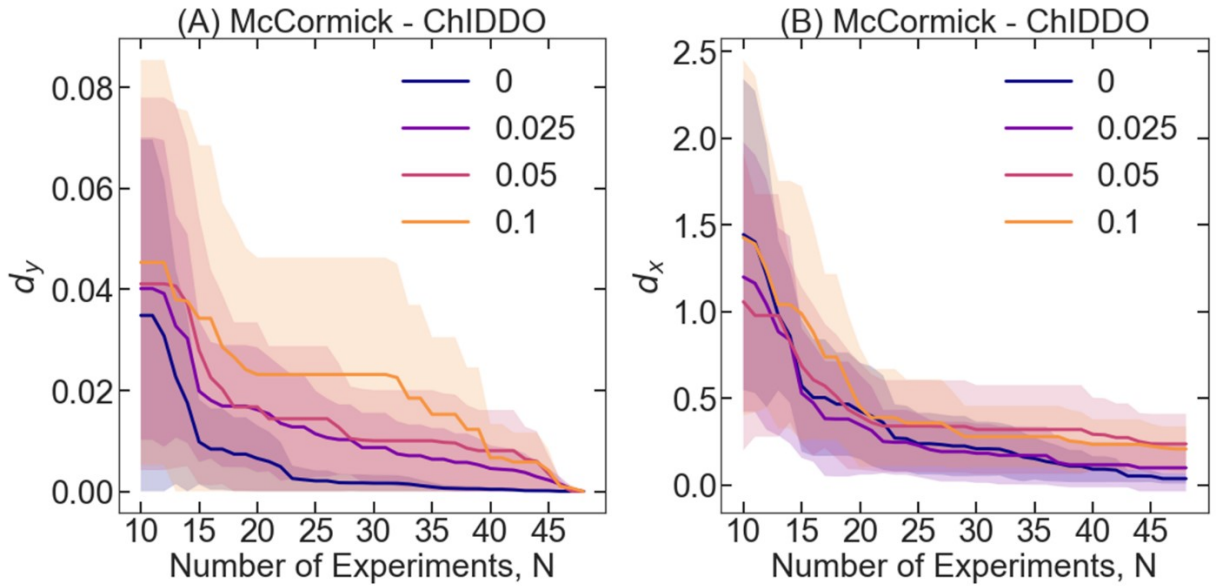


Figure S40. d_y and d_x versus number of experiments, N, comparing different noise levels using ChIDDO on the 2D Mccormick function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

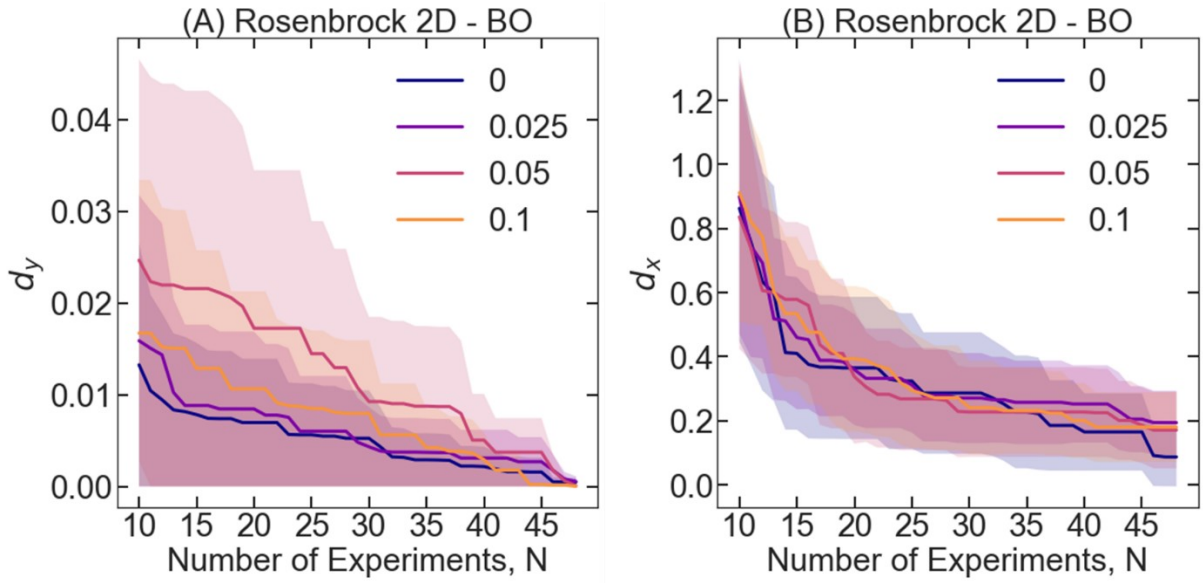


Figure S41. d_y and d_x versus number of experiments, N, comparing different noise levels using BO on the 2D Rosenbrock function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

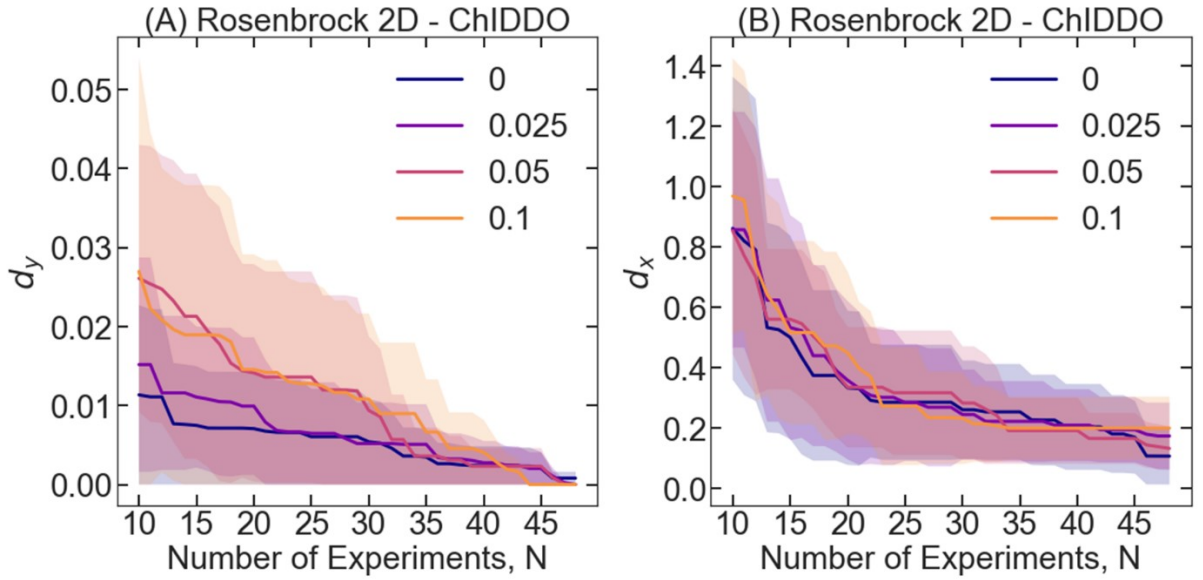


Figure S42. d_y and d_x versus number of experiments, N, comparing different noise levels using ChIDDO on the 2D Rosenbrock function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

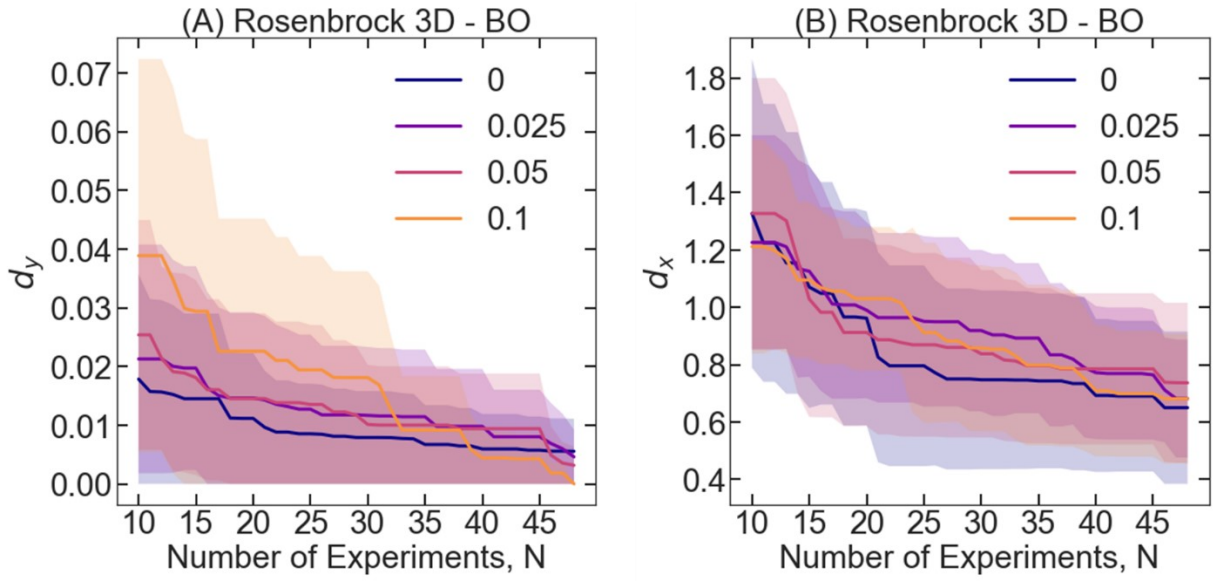


Figure S43. d_y and d_x versus number of experiments, N, comparing different noise levels using BO on the 3D Rosenbrock function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

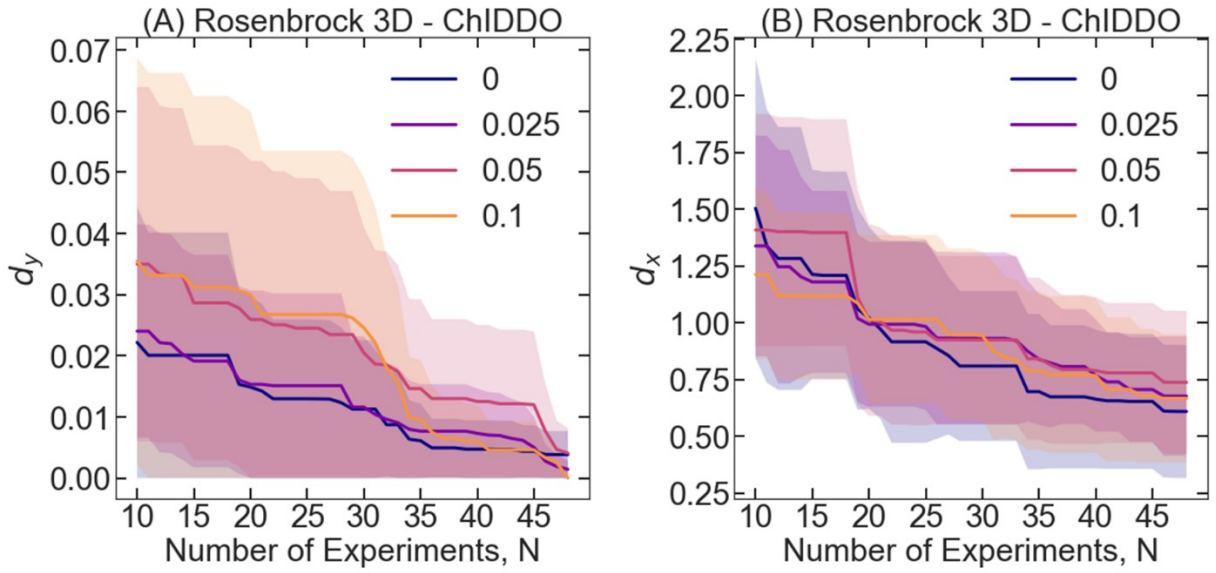


Figure S44. d_y and d_x versus number of experiments, N, comparing different noise levels using ChIDDO on the 3D Rosenbrock function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

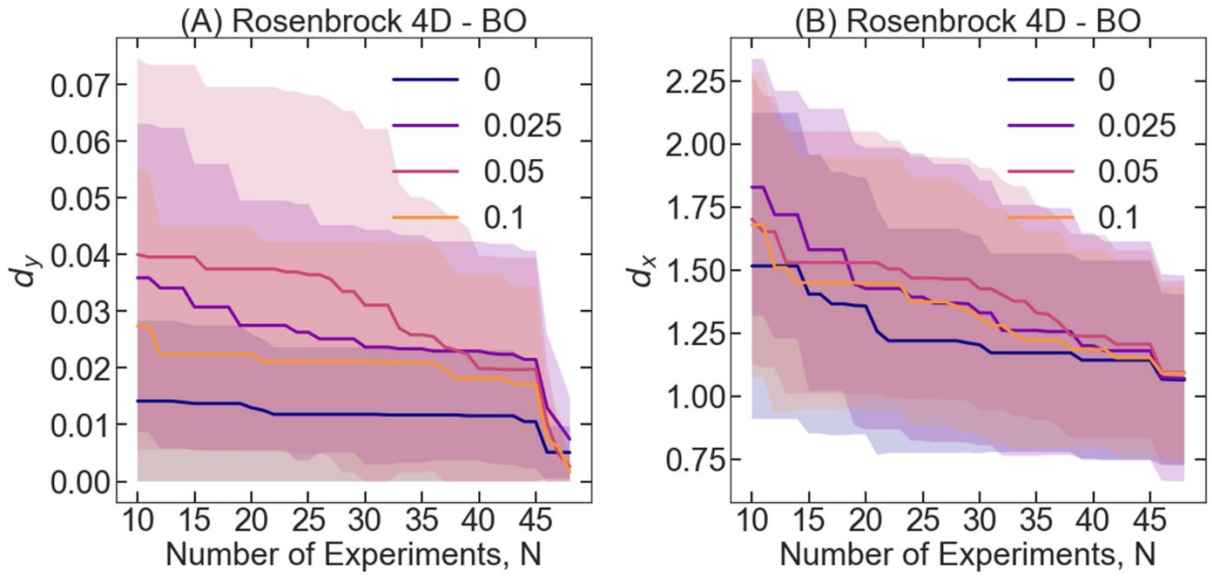


Figure S45. d_y and d_x versus number of experiments, N, comparing different noise levels using BO on the 4D Rosenbrock function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

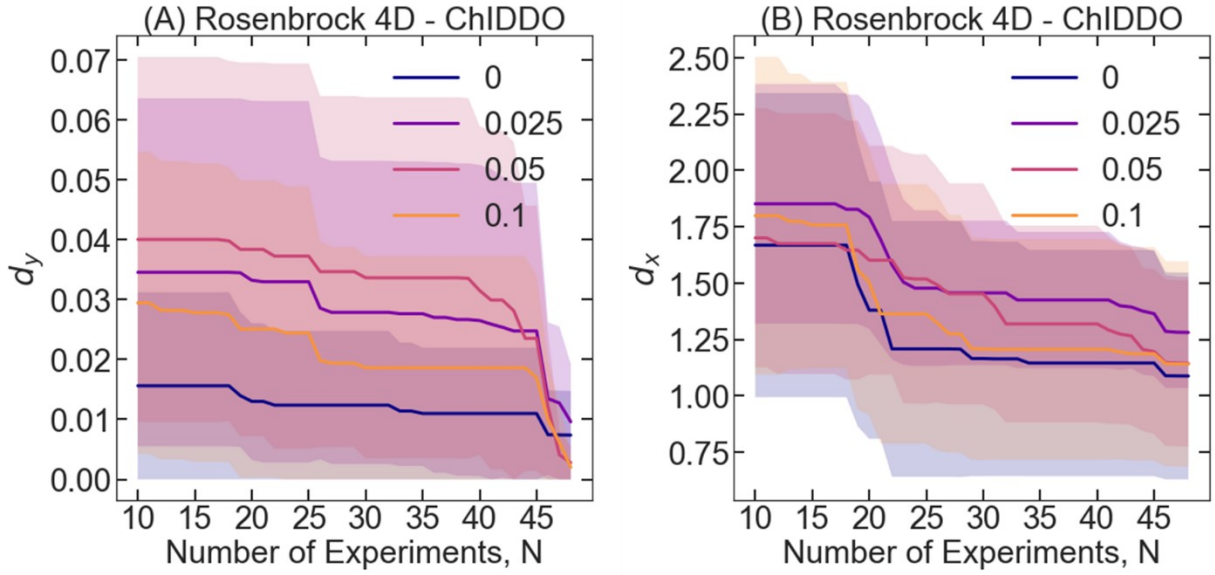


Figure S46. d_y and d_x versus number of experiments, N, comparing different noise levels using ChIDDO on the 4D Rosenbrock function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

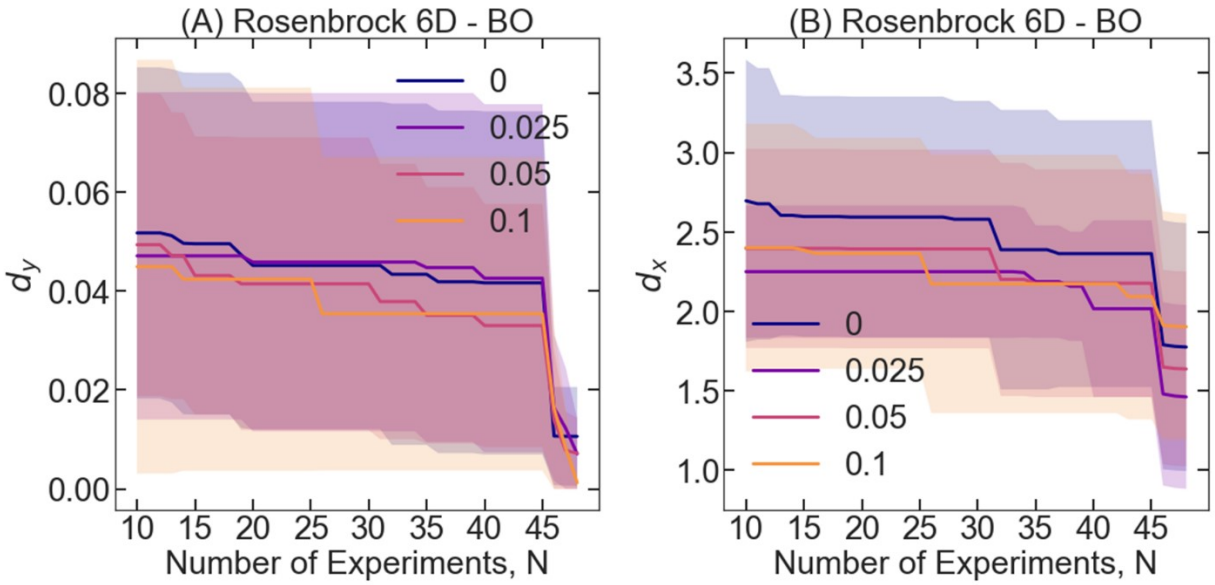


Figure S47. d_y and d_x versus number of experiments, N, comparing different noise levels using BO on the 6D Rosenbrock function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

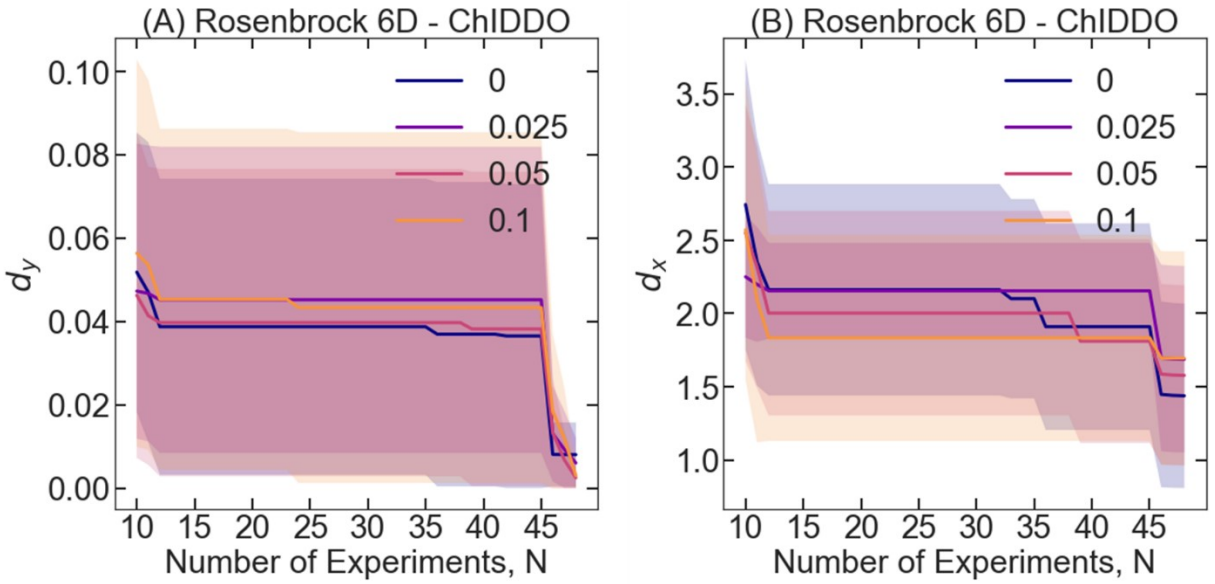


Figure S48. d_y and d_x versus number of experiments, N, comparing different noise levels using ChIDDO on the 6D Rosenbrock function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

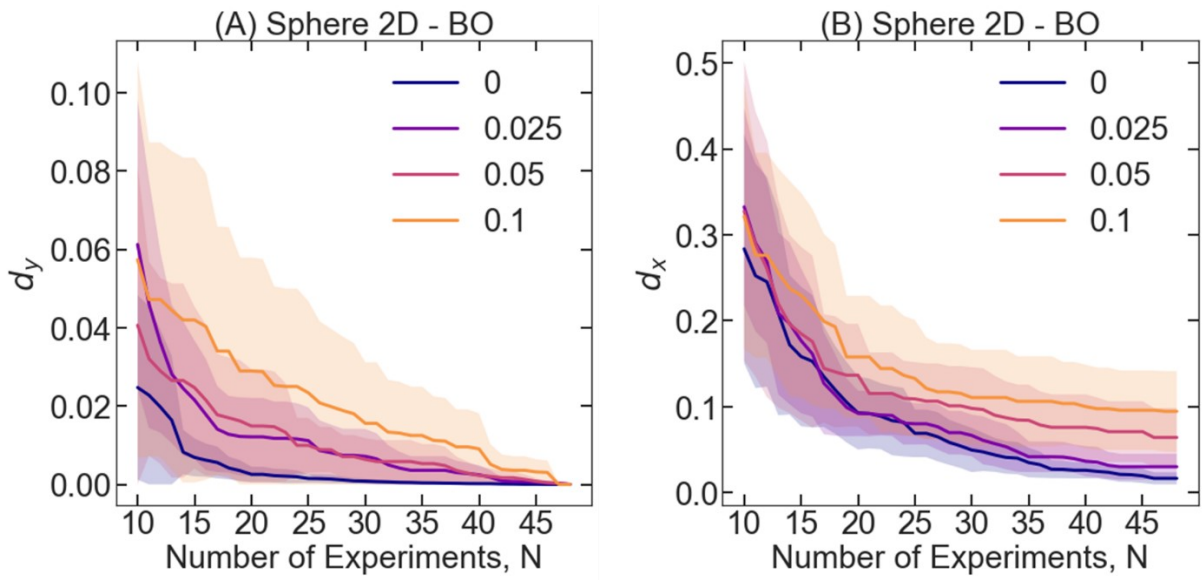


Figure S49. d_y and d_x versus number of experiments, N, comparing different noise levels using BO on the 2D Sphere function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

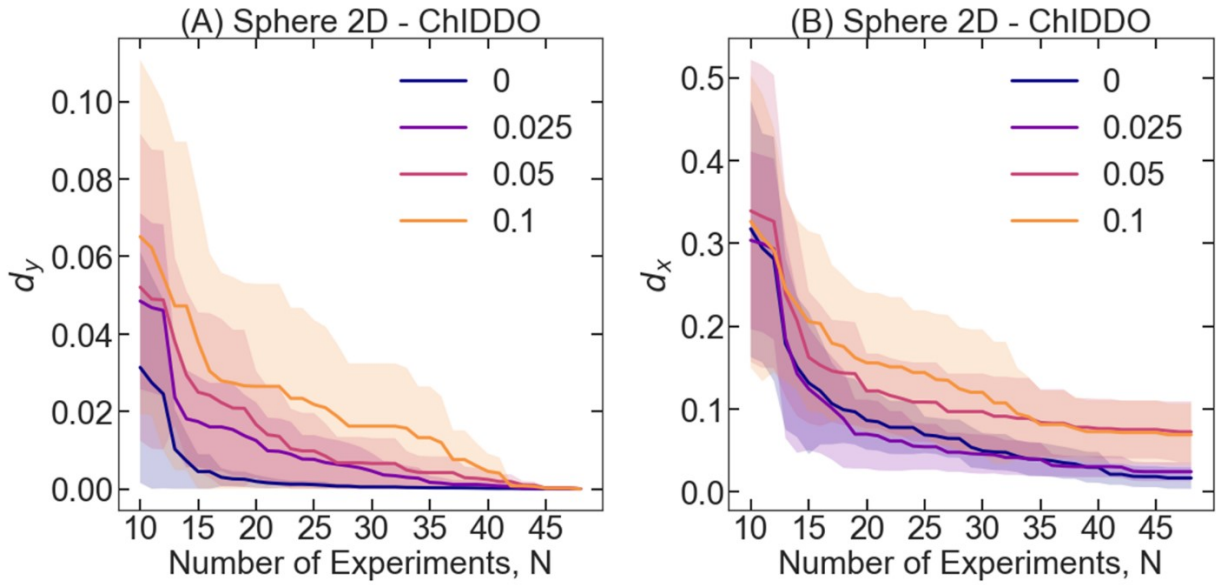


Figure S50. d_y and d_x versus number of experiments, N, comparing different noise levels using ChIDDO on the 2D Sphere function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

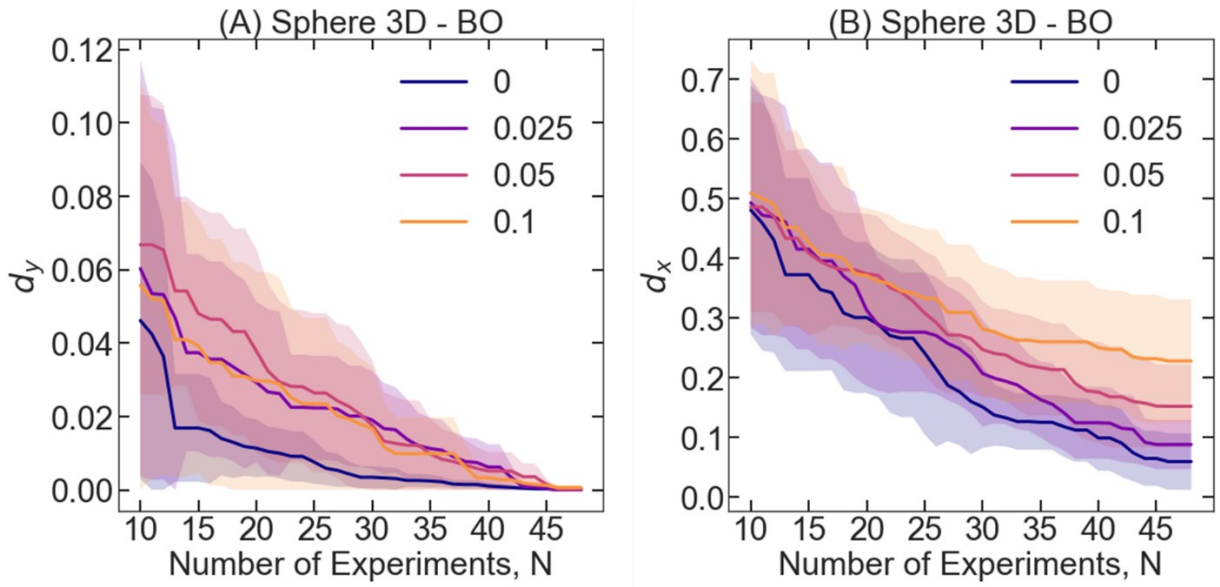


Figure S51. d_y and d_x versus number of experiments, N, comparing different noise levels using BO on the 3D Sphere function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

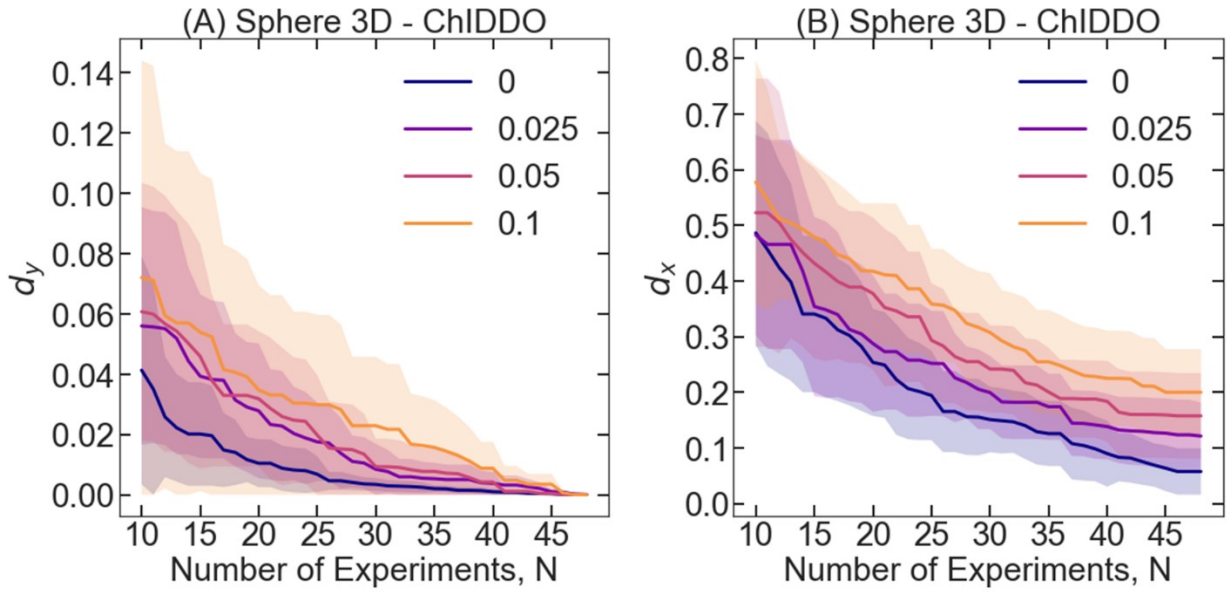


Figure S52. d_y and d_x versus number of experiments, N, comparing different noise levels using ChIDDO on the 3D Sphere function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

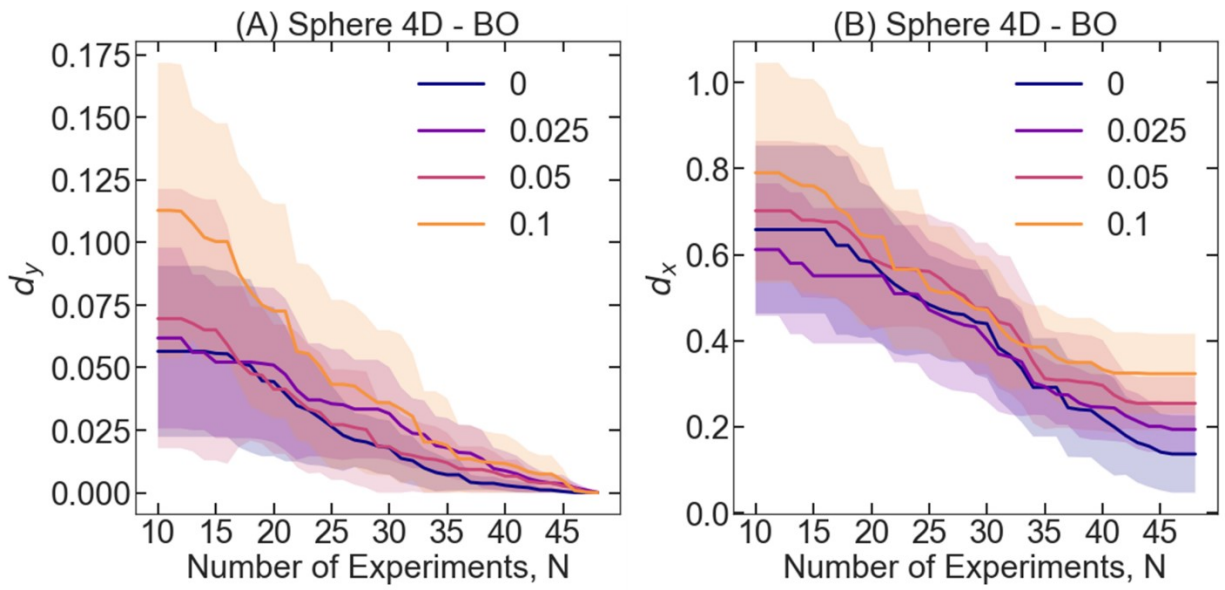


Figure S53. d_y and d_x versus number of experiments, N, comparing different noise levels using BO on the 4D Sphere function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

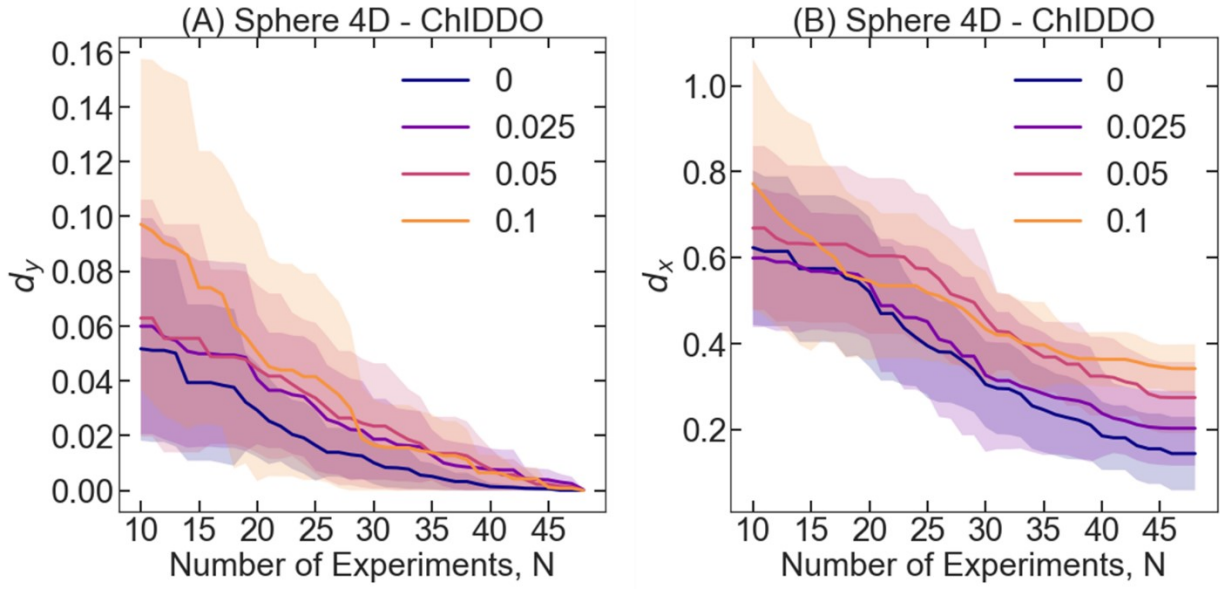


Figure S54. d_y and d_x versus number of experiments, N, comparing different noise levels using ChIDDO on the 4D Sphere function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

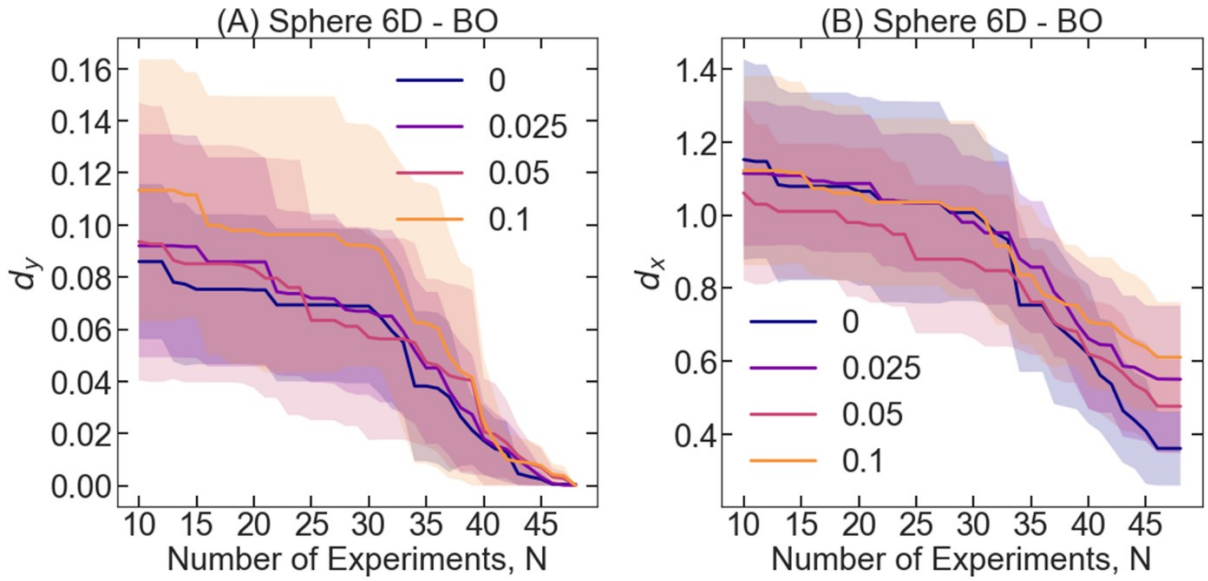


Figure S55. d_y and d_x versus number of experiments, N, comparing different noise levels using BO on the 6D Sphere function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

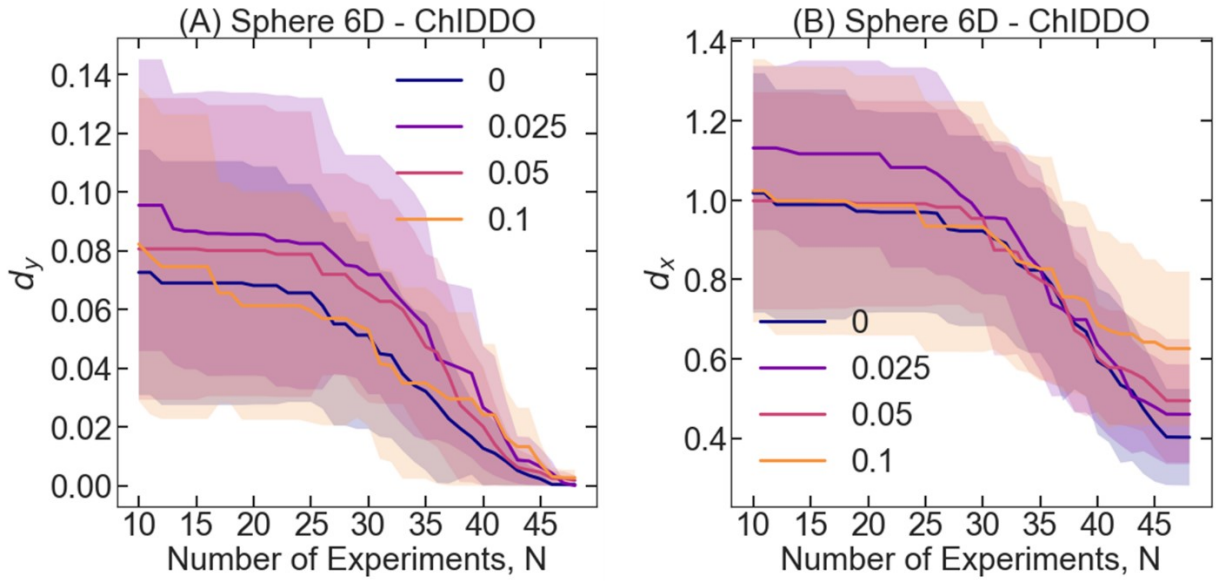


Figure S56. d_y and d_x versus number of experiments, N, comparing different noise levels using ChIDDO on the 6D Sphere function. (A) d_y, (B) d_x. For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. The MRB acquisition function was used.

Physics model accuracy comparison for other objective function combinations using MRB acquisition function

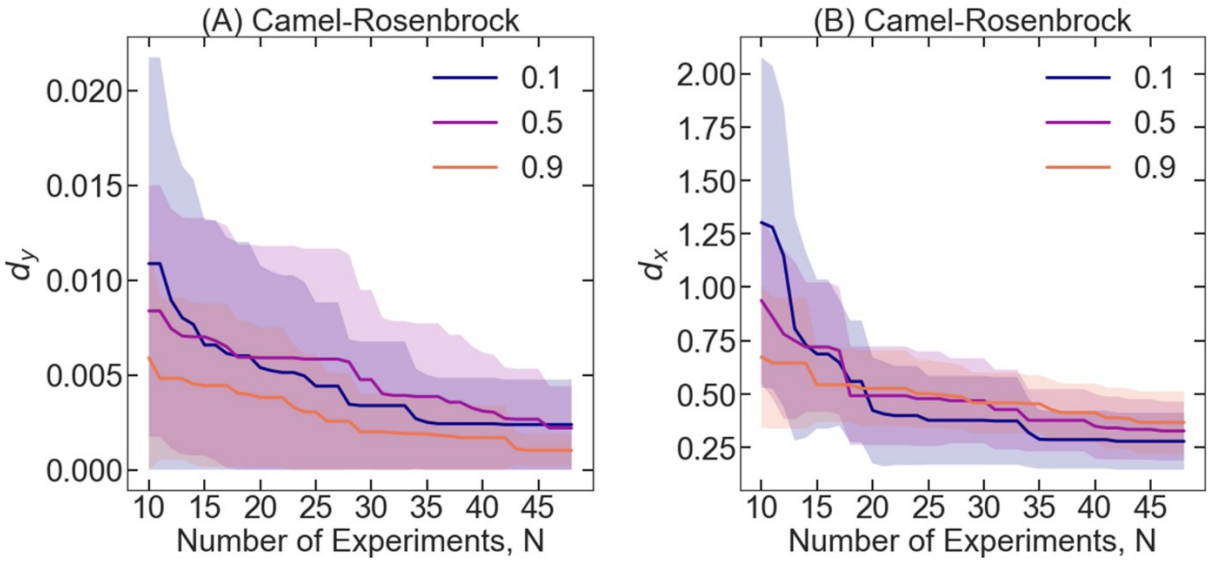


Figure S57. d_y and d_x versus number of experiments, N, for different objective function mixing ratios. 2D Camel mixed with 2D Rosenbrock using Camel as physics model. (A) d_y, (B) d_x .For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. For all of these graphs, ChIDDO was used as the AL algorithm and MRB was used as the acquisition function.

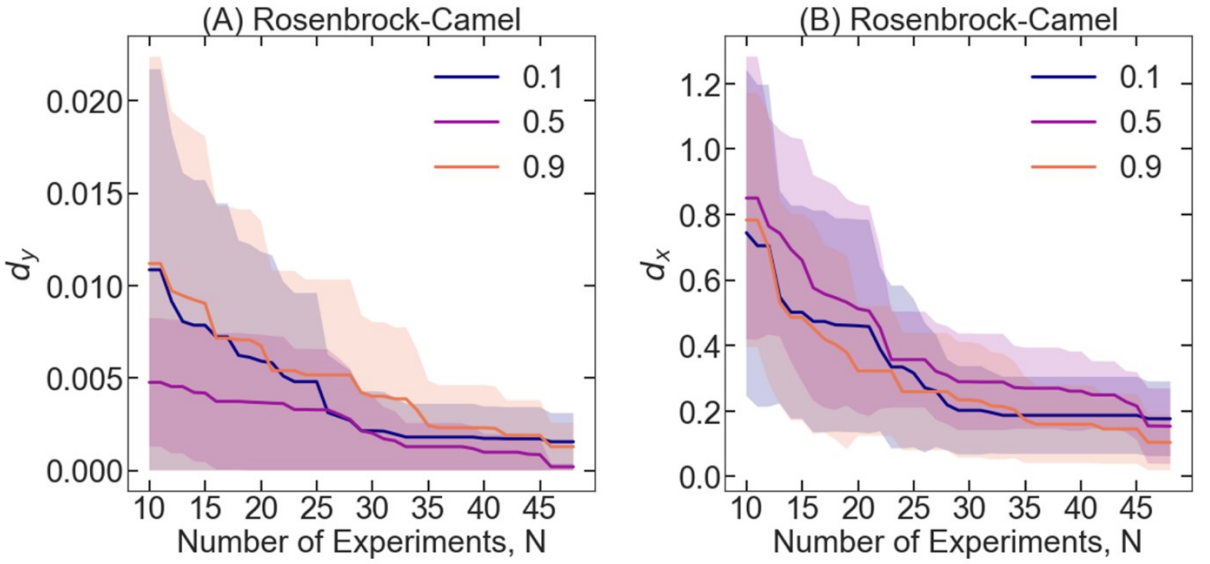


Figure S58. d_y and d_x versus number of experiments, N, for different objective function mixing ratios. 2D Camel mixed with 2D Rosenbrock using Rosenbrock as physics model. (A) d_y, (B) d_x .For each curve, 20 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. For all of these graphs, ChIDDO was used as the AL algorithm and MRB was used as the acquisition function.

Electrochemical simplified models using logarithmic updating of the exploration-exploitation rate

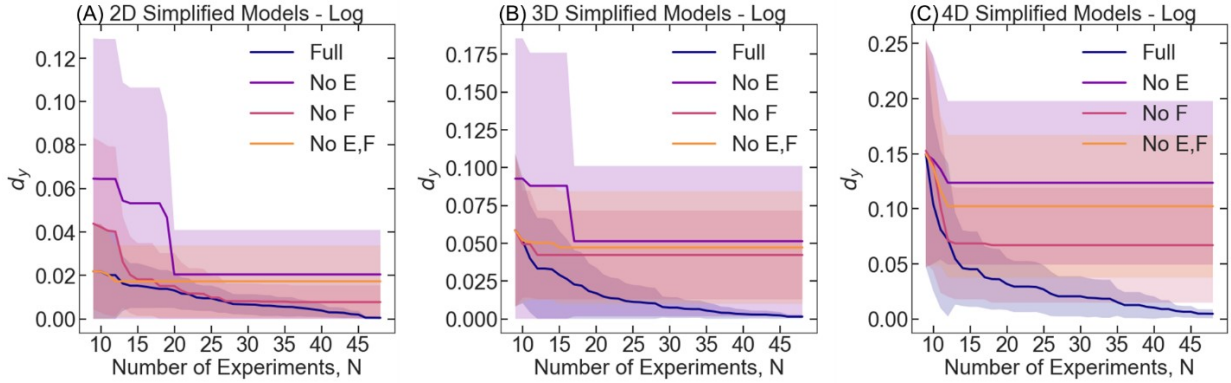


Figure S59. d_y versus number of experiments, N, for different electrochemical physics model information, using a logarithmic exploration/exploitation rate. "Full" indicates the model is predicting the same information as the objective function. "No E", "No F", and "No EF" indicate the removal of Equations 8 and/or 9 from the physics model information, resulting in a less informative model. (A) 2D electrochemical model. (B) 3D electrochemical model. (C) 4D electrochemical model. For each curve, 25 separate searches, S, were performed, and the average of the results are the lines shown. The shadow around each of the lines represents the standard deviation. For all of these graphs, ChIDDO was used as the AL algorithm and MRB was used as the acquisition function.

Supporting Information References

1. Acquisition funcitons. https://modal-python.readthedocs.io/en/latest/content/query_strategies/Acquisition-functions.html.

2. Poloczek, M.; Wang, J.; Frazier, P., Multi-information source optimization. *Advances in Neural Information Processing Systems* **2017**, 4288-4298.