Supplemental Material: Good Enough is Better: Feasibility vs. Pareto-Optimality in Alloy Design

Cayden Maguire[a], Christofer Hardcastle[a], Trevor Hastings[a], Raymundo Arróyave[a,b,c], Brent Vela[a,]

[a] *Department of Materials Science and Engineering, Texas A&M University, College Station, TX 77843, USA*

[b] *J. Mike Walker '66 Department of Mechanical Engineering, Texas A&M University, College Station, TX 77843, USA*

[c] *Wm Michael Barnes '64 Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX 77843, USA*

1. Constraint Satisfaction that Accounts for Objective Dependence

1.1. Methods

To account for dependence between objectives, we constructed a joint Gaussian distribution for the tasks (PoF Corr). Since each Gaussian process prediction is a normal random variable, a joint multivariate Gaussian distribution can be constructed if a positive semi-definite correlation matrix is defined over the objectives.

To estimate the correlation matrix, we computed Pearson product-moment correlation coefficients for each standardized GP residual for each queried alloy. Since we have four tasks, the correlation matrix is a 4×4 matrix. For a sequential Bayesian optimization scheme with one initial query, the correlation matrix cannot be computed until four iterations, when the dimension of the matrix equals the sample size (i.e., when the matrix is invertible). To account for this, we use the identity matrix for the initial three queries and only account for task correlations when the correlation matrix is invertible. We also applied a shrinkage factor to the correlation matrix. Shrinkage to the identity matrix avoids ill-conditioned correlation matrices, particularly when the number of samples is not large relative to the number of dimensions of the correlation matrix [1]. This is needed for our model, since the number of inputs will remain small compared to the dimension of the correlation matrix for many iterations. The matrix is of the form:

$$R_s = (1 - \lambda)R + \lambda I \tag{1}$$

where R_s is the correlation matrix with shrinkage applied, λ is the shrinkage factor (between 0 and 1, with a higher value shrinking the correlation matrix closer to the identity matrix), R is the correlation matrix computed using Pearson product-moment correlation coefficients, and I is the identity matrix. For our model, we chose a moderate shrinkage factor of 0.2. After constructing the correlation matrix, SciPy's multivariate normal cumulative density function (CDF) was used to compute the CDF and subsequently the probability of feasibility for each test point and each iteration.

1.2. Results and Discussion

Figure 1 shows the average cumulative feasible queries per iteration for the joint and indepdent PoF models, while Figure 2 shows the average increase in hypervolume. Both models were run using relaxed design constraints (see Section 2 below). For both cases, the joint PoF model performs similarly to the proposed (independent) PoF model, with both models having similar per iteration average metrics with overlapping standard deviations. The joint PoF model has a slightly higher uncertainty due to uncertainty propagation when joint mapping noisy marginal uncertainties. The aforementioned issues with estimating covariance for data sparse campaigns may also contribute to higher uncertainty [1].

It is also worth noting that the joint model is significantly more computationally expensive; the in-silico campaign took 23 hours and 55 minutes on our cluster, as opposed to the independent model, which took 43 minutes and 37 seconds using the same hardware (25 CPU cores with 240G of RAM). Monte Carlo sampling may be used to reduce the computational cost of the joint PoF model.

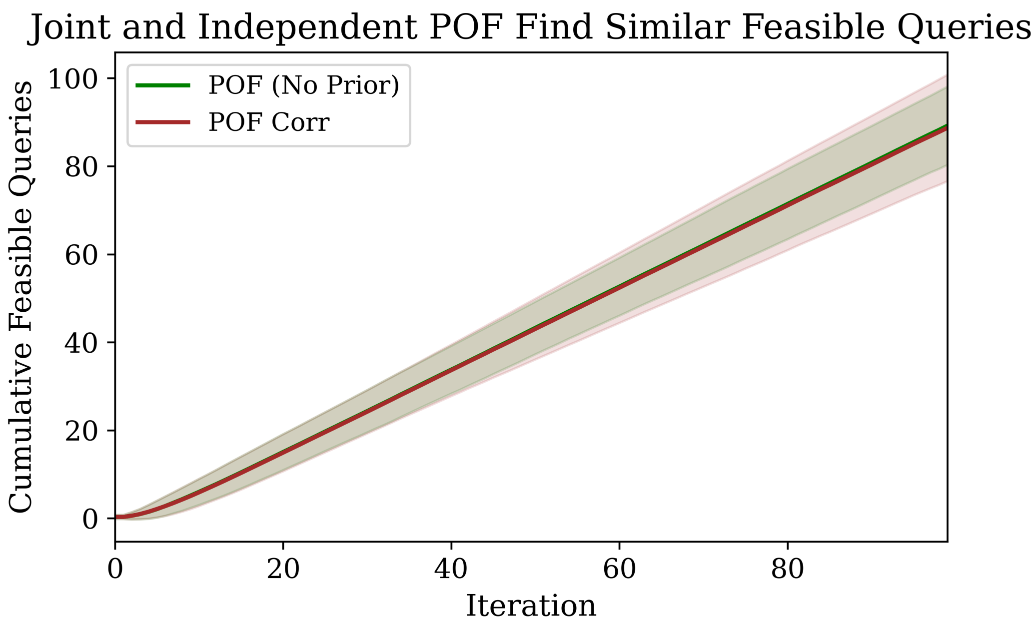

Figure 1: Average cumulative feasible queries per iteration for the independent PoF and joint PoF Corr models under relaxed constraints. Curves show mean performance across campaigns, with shaded regions indicating one standard deviation.

2. Results with Relaxed Design Constraints

All models were run with relaxed design constraints to assess generalizability. For this case, we set the constraints so that roughly 11% of the design space is feasible (4483 feasible alloys). These values are shown in Table 1. Figure 3 shows the average cumulative feasible query count per iteration while Figure 4 shows the average hypervolume increase per iteration.

As with the original run, the models with probability of feasibility (PoF) acquisition functions find more feasible queries by the end of the design campaign and find a feasible alloy earlier. On average, the PoF model (with and without a prior, and the model that accounts for objective dependence) find a feasible query by the fourth iteration, the combined model (with the acquisition function PEHVI PoF) finds a feasible query by the fifth iteration, and the optimization model and random search find a feasible query by the eighth iteration.

By the end of the campaign, the PoF models find approximately 90 feasible queries on average, nearly one feasible query per iteration for a 100 iteration sequential optimization scheme. The model with a prior results in a slightly higher average cumulative query count than the model without (91.83 vs 89.22 average cumulative queries). The joint

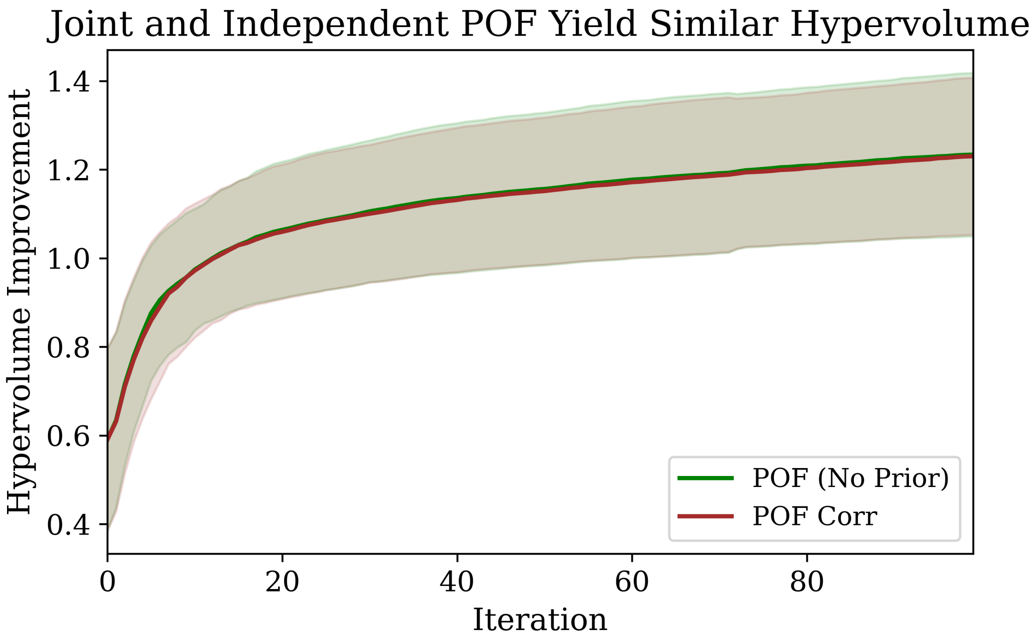

Figure 2: Average hypervolume improvement per iteration for the independent PoF and joint PoF Corr models under relaxed constraints. Curves show mean performance across campaigns, with shaded regions indicating one standard deviation.

PoF model (PoF Corr) performs similarly to the model that assumes dependence (88.73 vs 89.22 average cumulative queries).

Again, the models that use Pointwise Expected Hypervolume Improvement (PEHVI) result in a greater increase in hypervolume. The combined model performs in between the two strategies in terms of both hypervolume improvement and constraint satisfaction.

The model with the combined acquisition function performs worse at constraint satisfaction than the models that only consider PoF, as the solution space is limited to candidates that are also expected to increase hypervolume. Candidates that maximize hypervolume are not guaranteed to be feasible, as evident by the random search resulting in a greater hypervolume increase but less feasible queries than the PoF models. Hence, a combined strategy may miss candidates that are feasible, but result in a more modest hypervolume increase. For highly constrained problems, PoF is the better option.

Table 1: Feasibility constraints

Property	Criterion	Threshold	Notes
Density	$<$	13.15 g/cm^3	70^{th} percentile of dataset
Yield Strength at 600°C	$>$	871.75 MPa	30^{th} percentile of dataset
Pugh Ratio	$>$	2.218	30^{th} percentile of dataset
Solidus Temperature	$>$	2252 K	Fixed threshold
BCC phase flag at 600°C	=	5.0	Single-phase BCC indicator

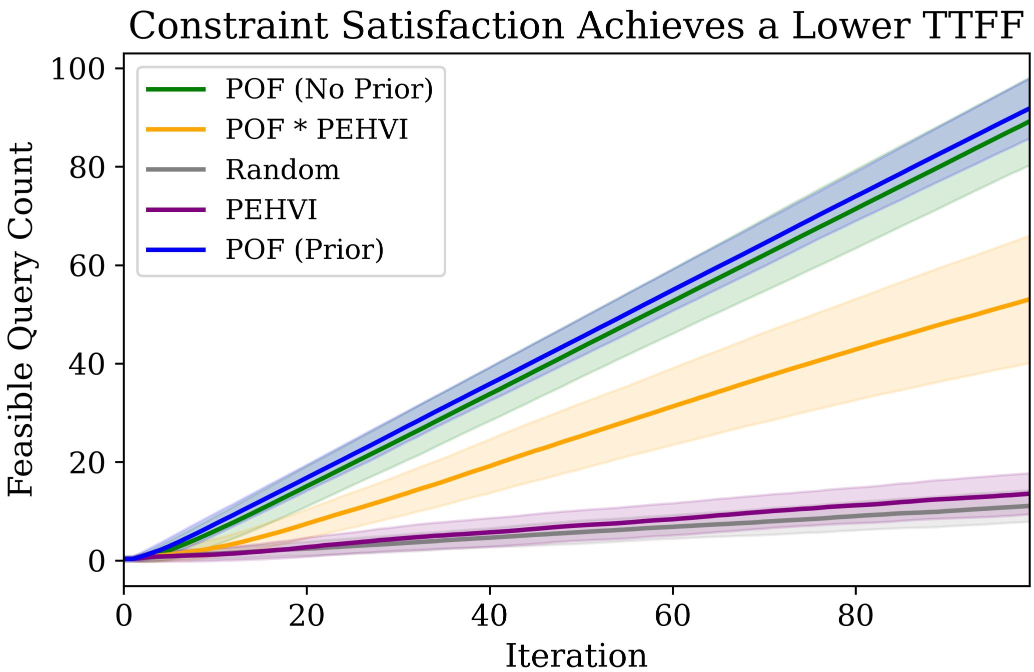

Figure 3: Average cumulative feasible queries per iteration for all acquisition strategies evaluated with relaxed constraints. Probability-of-feasibility variants identify feasible candidates earlier and maintain higher feasible-query counts over the campaign.

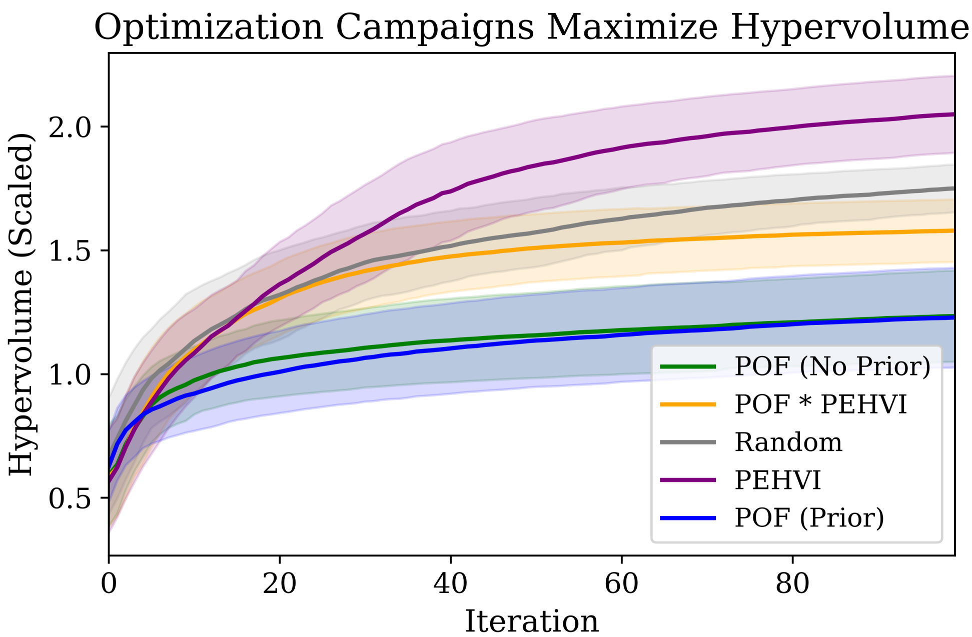

Figure 4: Average hypervolume improvement per iteration for all acquisition strategies evaluated with relaxed constraints. PEHVI-based strategies produce larger hypervolume gains, while PoF-focused strategies prioritize feasibility discovery.

References

[1] O. Ledoit, M. Wolf, A well-conditioned estimator for large-dimensional covariance matrices, Journal of Multivariate Analysis 88 (2) (2004) 365–411. doi:https://doi.org/10.1016/S0047-259X(03)00096-4.

URL https://www.sciencedirect.com/science/article/pii/S0047259X03000964