Electronic Supplementary Material

Using machine learning models to explore the solution space of large nonlinear systems underlying flowsheet simulations with constraints

Patrick Otto Ludl (✉)[1], Raoul Heese[1], Johannes Höller[1],

Norbert Asprion[2], Michael Bortz[1]

1 Fraunhofer ITWM Optimization Department, Kaiserslautern 67663, Germany
2 Chemical and Process Engineering BASF SE, Ludwigshafen 67056, Germany

E-mail: patrick.otto.ludl@itwm.fraunhofer.de

Appendix A Explicit inequality constraint probability

In this appendix section, we briefly show how the integral from Eq. (13) can be solved explicitly for the special case of (i) inequality constraints

$$f_{i,\min} \le f_i \le f_{i,\max} \quad \forall i = 1, \dots, p, \tag{A1}$$

i. e.,

$$D_f = [f_{1,\min}, f_{1,\max}] \times \cdots \times [f_{p,\min}, f_{p,\max}] \tag{A2}$$

with $f = (f_1, \dots, f_p)$ and (ii) a multivariate Gaussian distribution

$$\pi_f(f|x) = \frac{1}{(2\pi)^{\frac{p}{2}}} \frac{1}{(\det\Sigma)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(f-\bar{f})^T \Sigma^{-1} (f-\bar{f})\right\}. \tag{A3}$$

The probability $p(f \in D_f|x)$ is then given by

$$p\left(f \in D_f \middle| x\right) = \prod_{i=1}^{p} \int_{f_{i,\min}}^{f_{i,\max}} \pi_f\,(f|x)\mathrm{d}f_i. \tag{A4}$$

Diagonalizing the covariance matrix as $\Sigma = U^T \mathrm{diag}(\sigma_1^2, \dots, \sigma_p^2) U$ with an orthogonal $p \times p$-matrix U and changing the integration variables to

$$\tilde{f} = U\left(f - \bar{f}\right) \tag{A5}$$

leads to

$$p\left(f \in D_f \middle| x\right) = \prod_{i=1}^{p} \int_{\tilde{f}_{i,\min}}^{\tilde{f}_{i,\max}} \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{\tilde{f}_i^2}{2\sigma_i^2}} \mathrm{d}\tilde{f}_i, \tag{A6}$$

where

$$\tilde{f}_{\min,\max} = U\left(f_{\min,\max} - \bar{f}\right) \tag{A7}$$

and we symbolically use $\mathrm{d}\tilde{f} = |\det U| \mathrm{d}f = \mathrm{d}f$. The integrals in Eq. (A6) can be expressed by means of the error function

$$\mathrm{erf}(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^{x} e^{-t^2}\, \mathrm{d}t = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-t^2}\, \mathrm{d}t \tag{A8}$$

to yield

$$p\left(f \in D_f \middle| x\right) = \frac{1}{2^p} \prod_{i=1}^{p} \left\{ \mathrm{erf}\left(\frac{\tilde{f}_{i,\max}}{\sqrt{2\sigma_i^2}} \right) - \mathrm{erf}\left(\frac{\tilde{f}_{i,\min}}{\sqrt{2\sigma_i^2}} \right) \right\}. \tag{A9}$$

This explicit form reduces the computational effort of computing $p(f \in D_f|x)$. In all examples presented in this paper GPR is used as the regression model $\mathcal{R}_f$ and all constraints are inequality constraints. Thus, Eq. (A9) is applied in all examples.

B Incorporating inequality constraints directly in the optimization problem

We have included the preference for the fulfillment of constraints as an additional utility term U_c in Algorithm 2. In many cases, the constraints imposed on the process, Eq. (7), are given in the form of inequality constraints as shown in Eqs. (A1) and (A2), respectively. This allows the construction of an alternative adaptive sampling algorithm by omitting U_c in the utility function (i.e., $w_c = 0$) and instead including the inequality constraints directly into the optimization problem, Eq. (3).

Specifically, the regression model $\mathcal{R}_f$ predicts the constraint function values f, i.e.,

$$\mathcal{R}_f : x \mapsto \hat{f}(x) \tag{A10}$$

and therefore allows to translate the inequality constraints, Eq. (A1), to constraints on x according to

$$f_{i,\min} \leq \hat{f}_i(x) \leq f_{i,\max} \quad \forall i = 1, \dots, p. \tag{A11}$$

Since $\mathcal{R}_f$ is only an approximation of the true function $f(x)$, the inequality constraints are also only approximations of the true constraints. However, they can be formally used to define the constrained optimization problem

$$x_{\text{new}} = \operatorname{argmax}_{x \in \chi,\, \hat{f}(x) \in D_f} U\left(D_{\text{expl}}, x, w\right)|_{w_c = 0} \tag{A12}$$

as a replacement for Eq. (3). Note that the utility function with $w_c = 0$ does not depend on $\mathcal{R}_f$.

Summarized, by approximating the constraints, Eq. (A1), by Eq. (A11) and integrating them directly into the optimization problem of the utility function without using U_c, Eq. (A12), we obtain Algorithm 3. The drawback of this way of including constraints in the adaptive sampling strategy is the lack of possibility to set the weight of the constraints and the requirement of a sufficiently good surrogate model $\mathcal{R}_f$ to get a meaningful prediction of the constraint fulfillment.

Algorithm 3 Alternative to Algorithm 2: We include inequality constraints $f \in D_f$ directly into the optimization of the utility function. The new function SuggestionConstrainedDirect replaces SuggestionUnconstrained in Line 5 of Algorithm 1.

function SuggestionConstrainedDirect($D_{\text{expl}}, \chi, D_f, w$)
    $\mathcal{C} \leftarrow$ TrainClassifier(D_{expl})
    $\mathcal{R}_f \leftarrow$ TrainRegressor(D_{expl})
    function Utility($x, D_{\text{expl}}, \mathcal{C}, w$)
        $u \leftarrow (U_s(\mathcal{C}, x), U_r'(D_{\text{expl}}, x))^T$
        return $\frac{w^T u}{||w||_1}$
    end function
    return $\operatorname{argmax}_{x \in \chi,\, \hat{f}(x) \in D_f}$ Utility$(x, D_{\text{expl}}, \mathcal{C}, w)|_{w_c=0}$
end function