LETTERS TO THE EDITOR

To the Editor:

In the paper titled "Diagnostic Model Processor: Using Deep Knowledge for Process Fault Diagnosis" (April 1990, p. 565), Petti, Klein, and Dhurjati presented a method of malfunction diagnosis based on the violation/satisfaction of model constraint equations. A critical formula in this work is the one for evaluating evidence and rating malfunctions (Eq. 4). Since this formula only measures the extent to which the expected abnormal symptoms of each potential fault are displayed, it does not help determine how many faults or what combinations of multiple faults might be present. A much more focused interpretation is readily derived from the same constraint violation/satisfaction data used by these authors.

Consider a case where there are four measured constraints, C_1-C_4, that are satisfied unless certain faults a_i (representing malfunctions or violation of modeling assumptions) are present. The faults need not be mutually exclusive. Assume the effects of the faults on the constraints:

	C_1	C_2	C_3	C_4
a_1	x	0	0	0
a_2	0	x	0	0
a_3	x	x	x	0
a_4	0	0	x	0
a_5	x	x	x	x

x = constraint violated
0 = no effect of fault on constraint

Equation 4 of Petti et al. produces a "likelihood" for each fault given a pattern of satisfied and violated constraints. Suppose, in this case, constraints C_1, C_2 and C_3 are violated, and C_4 is satisfied. Assuming unit sensitivities for all entries, Eq. 4 yields:

$$\text{Likelihood}(a_1) = (1*1 + 0*1 + 0*1 + 0*0)/1 = 1$$

$$\text{Likelihood}(a_2) = (0*1 + 1*1 + 0*1 + 0*0)/1 = 1$$

$$\text{Likelihood}(a_3) = (1*1 + 1*1 + 1*1 + 0*0)/3 = 1$$

$$\text{Likelihood}(a_4) = (0*1 + 0*1 + 1*1 + 0*0)/1 = 1$$

$$\text{Likelihood}(a_5) = (1*1 + 1*1 + 1*1 + 0*1)/4 = 0.75$$

This formula, therefore, predicts that faults 1–4 are equally "likely," since the symptoms of each fault are fully displayed, with a_5 a lesser but still likely candidate. Besides eliminating no possibilities, the proposed likelihoods give no information concerning how many faults are actually present and in what combinations. For example, from the likelihoods one cannot tell that a_3 by itself is adequate to explain the violated constraints, while faults a_1 and a_2 individually are not.

It is straightforward to correct this situation. Examination of the fault/constraint table shows that the possible states for the plant (temporarily neglecting uncertainties) are:

H_1: a_3
H_2: a_1 & a_2 & a_4
H_3: a_1 & a_3
H_4: a_2 & a_3
H_5: a_4 & a_3
H_6: a_1 & a_2 & a_3
.
.
.
H_9: a_1 & a_2 & a_3 & a_4

Since $Pr(A \& B) \leq Pr(A)$, only H_1 and H_2 are noninferior, minimal spanning sets; the combinations of faults H_3 – H_9 are less probable. If faults are independent and prior probabilities are roughly equal, the single fault hypothesis H_1 is most likely, but if there are unequal prior probabilities and dependencies among the faults, it is possible that H_2 is the best candidate.

Finding single- and multiple-fault explanations is identical to finding minimal cut sets in fault trees. For the current example, the fault tree would be drawn as Figure 1. The satisfied constraint C_4 is indicated by the negation. Evaluation of the minimal cut sets of this tree yields the two noninferior hypotheses, H_1 and H_2. After determination of the minimal cut sets based on the observed qualitative state of the plant, evidence weighting formulas can be used to check the goodness of fit between the hypothesis and the observed constraint satisfactions/violations. The formula of Petti et al. is suggestive of a ranking based on the colinearity of actual and predicted constraint patterns. What follows is only one among many conceivable schemes for ranking

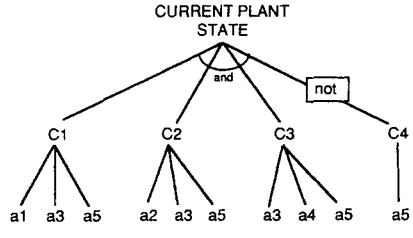

Figure 1. Fault tree.

hypotheses, but it does have the capability of ranking *any* single- or multiple-fault hypotheses in a reasonable order, whether or not they are generated from fault tree analysis.

Let ϵ be the vector of constraint residuals, and S be the fault/constraint sensitivity matrix, S_{ij} the sensitivity of constraint j to fault i, without division by τ as in the original paper. Let s be a matrix whose columns are the rows of S corresponding to the faults in the hypothesis of interest. Let f be a column vector of fault extents, whose length corresponds to the number of faults in the current hypothesis. The residuals should satisfy:

$$\epsilon = sf + v$$

where v represents the contribution of noise to the residuals, with covariance R. The optimum value for f can be determined by least squares or maximum likelihood estimation:

$$f^* = (s^t R^{-1} s)^{-1} s^t R^{-1} \epsilon$$

where $R = I$ for least squares. Let ϵ^* be the predicted constraint residual under the current hypothesis:

$$\epsilon^* = sf^*$$

A measure of sufficiency of the hypothesis could then be the colinearity of the vectors ϵ and ϵ^*:

$$F(H_i) = \frac{\epsilon^t \epsilon^*}{\|\epsilon\| \cdot \|\epsilon^*\|}$$

$F(H_i)$ represents the likelihood of hypothesis H_i before normalization or application of prior probabilities. The final ranking of each fault hypothesis is the product of the calculated F value and the (estimated) prior probability of the hypothesis. Since in general the prior probabilities for hypotheses involving multiple faults are less than for single fault hypotheses, this method will tend to favor simpler explanations over more complex ones. Applying this formula to the noninferior hypotheses, assuming $R = I$,

$$f^*(H_1) = 1 \qquad \Rightarrow F(H_1) = 1$$
$$f^*(H_2) = (1\ 1\ 1)^t \Rightarrow F(H_2) = 1$$

The inferior hypotheses $H_3 - H_9$ will also have $F = 1$, since they can perfectly fit the observed pattern of residuals, but the prior probabilities are lower for the nonminimal hypotheses. Therefore, the final rankings of the nonminimal hypotheses $H_3 - H_9$ will be lower than for H_1 and/or H_2. This technique can also calculate likelihoods for other single fault hypotheses:

$$F(a_1) = 1/\sqrt{3} \approx 0.58$$
$$F(a_2) = 1/\sqrt{3}$$
$$F(a_4) = 1/\sqrt{3}$$
$$F(a_5) = \sqrt{3}/2 \approx 0.87$$

This shows that a_1 and a_2 are not sufficient in and of themselves to describe the pattern of constraint violations and satisfactions, which was not apparent through Petti et al.'s likelihoods. Furthermore, the result conforms to our intuition that a_5 is a better single-fault hypothesis than a_1, a_2, or a_4 since it fails to match only one constraint rather than two out of the four constraint pattern. This basic formulation can be adapted if there are complications such as constraints on f or uncertainties in the sensitivity coefficients.

An unrelated point. The formula used by Petti et al. to calculate tolerances on derived constraints given in Eq. 6 will give too large a tolerance. Tolerances represent statistical limits on equations and should be added in a manner analogous to the sum of variances: $\tau(z)^2 = \tau(x)^2 + c^2\tau(y)^2$, where z is the secondary equation given by $z = x + cy$. The resulting tolerance is always smaller than that given by Petti et al., by the triangle inequality.

Mark A. Kramer
Department of Chemical Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139

Reply:

We would like to thank Dr. Kramer for his interest in this work and for taking the time to examine the implications of the diagnostic model processor. This methodology, which is described in Petti et al. (1990), incorporates important features such as non-Boolean reasoning, use of an explicit model, consideration of equation sensitivities, and the possibility of multiple fault explanations. It is clearly stated in the paper that the method often generates multiple fault possibilities which may or may not be occurring, and under these conditions a further layer of intelligence may be necessary to refine the diagnosis. We have taken the approach of not making any assumptions to prune the possible fault set at this stage so that during refinement any simplifications which are made may be explicitly stated while the complete fault set from the diagnostic model processor remains available. Dr. Kramer's proposed changes violate this philosophy by limiting the possible hypotheses based on a Boolean pattern of failed constraints before evaluation of the evidence. Additionally, his evidence evaluation method does not succeed in properly evaluating all hypotheses.

Dr. Kramer's major criticism of the diagnostic model processor lies in the evidence evaluation formula (Eq. 4). He states that this is not a good approach because it "does not help to determine how many faults or what combinations of multiple faults are present." In his simple example, he shows that the diagnostic model processor produces equal likelihood of failure for many of the fault possibilities. While the results do not yield much information, given the poorly formulated constraint set, they are the best that can be obtained at this stage of the diagnosis.

Additionally, unless we rely on some method of hypothesis generation and thereby prune our allowable faults, this is the only possible result. For the example in Dr. Kramer's letter, he uses a Boolean decision tree to generate a hypothesis set which is later evaluated using an estimation technique. This straightforward decision tree method is not generally applicable in a non-Boolean approach for two reasons.

1. The constraint equations are not

categorized as "violated" and "satisfied," but are handled with a measure of their satisfaction.

2. The fault/constraint matrix contains sensitivities which describe the relationship between each equation and the possible faults, not just x's which indicate dependence.

Due to these two points, it is unclear what combinations of degree of failure and associated sensitivities are sufficient to include or eliminate the consideration of a hypothesis. The diagnostic model processor eliminates the need for generating fault hypotheses prior to evidence testing by assuming that all constraints have failed and that all faults are present to some degree. In taking this approach we ensure that fault possibilities are not inadvertently discounted.

In the example in Dr. Kramer's letter, we can assume unit sensitivities to generate the comparison matrix as described in Petti et al. (Eq. 10) and shown in Table 1.

Table 1.

	$C =$				
	a_1	a_2	a_3	a_4	a_5
a_1	—	0	1/3	0	1/4
a_2	0	—	1/3	0	1/4
a_3	1	1	—	1	3/4
a_4	0	0	1/3	—	1/4
a_5	1	1	1	1	—

It is obvious that the expected performance of the diagnostic model processor from these constraint relations will result in some inconclusive diagnoses (many 1's). Examination of this matrix suggests the need for additional constraint equations to help improve the discrimination among the assumptions. Ideally, the equations used for diagnosis of one assumption will never be an exact subset of the equations used in the diagnosis of another.

This is certainly not the case in this example. If no other constraints can be formulated, however, this matrix indicates the areas of expected problems and refinement techniques can be directed to resolve these discrimination problems. Examples of refinement techniques are the use of a single-fault hypothesis, use of known prior probabilities of equipment failure, consideration to the relationship between possible faults and the use of heuristic information. The advantage of this approach is that these techniques need be used only if necessary, and when they are used, the results are taken knowing that a simplification has been made to narrow the fault set. In the event that these refinement techniques generate incorrect or inconclusive results, the original fault set from the diagnostic model processor is always available. The implication is that no information is thrown away in an early stage of the diagnosis.

Assuming some method of hypothesis generation has been used, Dr. Kramer proposes a method of testing which is designed to rank the various hypotheses. He suggests using an estimation algorithm to fit a vector of fault extents to the observed residuals, with the claim that it has "the capability of ranking *any* single- or multiple-fault hypotheses in a reasonable order." This claim is not true because the method will always fit the observed residuals better when the order of the parameter vector is increased. This corresponds to the result of $F(a_i)$ always being less than or equal to $F(a_i \text{ and } a_j)$. The additional member, a_j, of the fault extent vector, f can always be set to zero. This reduces the estimation to the same problem as considering a_i only, therefore guaranteeing that the fit be no worse. This is exactly the result which Dr. Kramer obtains when testing his "inferior" hypotheses; they all produce an $F = 1$. Therefore in the final ranking, the multiple fault hypotheses will be judged solely on the estimated prior probabilities, regardless of whether the additional fault member, a_j, improves the fit of the evidence. This problem can be shown to be more critical if the hypothesis is considered which includes the failure of assumption a_5. Although this assumption is excluded in the Boolean decision tree, if all hypotheses are tested then the evidence evaluation formula should be able to rank these hypotheses as being less likely. However, Dr. Kramer's approach for a hypothesis such as: a_1, a_2, a_4 and a_5 are all failing, also produces a $F(a_1, a_2, a_4 \text{ and } a_5) = 1$; similarly, $F(a_3 \text{ and } a_5) = F(a_1, a_3 \text{ and } a_5) = F(a_2, a_3 \text{ and } a_5) = F(a_3, a_4 \text{ and } a_5) = 1$. The final ranking depends only on the prior probabilities. For example, compare the ranking for the fault hypothesis: a_3 and a_1, a "possible state" from the decision tree, and the fault hypothesis: a_3 and a_5, an impossible state from the decision tree. If the estimated prior probabilities of event a_1 happened to be less than that of event a_5, then the ranking of the hypothesis: a_3 and a_5 would be higher than that of hypothesis: a_3 and a_1. The method, therefore, fails to rank these hypotheses in a reasonable order. It is clear that this approach is dependent on the method used for hypothesis generation and the pruning of the hypothesis set based on prior probabilities. No new information has been found by using this fitting technique; all the work had been done when the most probable hypotheses were specified and when the prior probabilities were estimated. As has been discussed this is not reasonable in a true non-Boolean approach and the diagnostic model processor imposes no such limitations.

Additional problems exist with this approach due to the nature of estimation methods for determining an optimal f. One problem will always exist when the number of assumptions in a hypothesis is greater than that of constraint equations in the system. Under these conditions, the estimation problem is underdetermined and no unique solution exists so that the resulting f will be able to fit the residuals regardless of the validity of the hypothesis. The estimation technique will not be capable of handling these cases however, because the resulting observation matrix, s, will not be maximum rank. An example of this type of problem is Dr. Kramer's H_9, where a_1, a_2, a_3, and a_4 are all failed, although the problem is not underspecified, these assumptions are linearly-dependent. He claims that the estimation technique will produce an $F(H_9) = 1$; however, when we examine the matrix s for H_9:

$$s = \begin{matrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{matrix}$$

It is clearly not maximum rank. This implies that the matrix ($s's$) will be singular and the estimation technique will not work.

Although problems exist in discriminating multiple-fault hypotheses with Dr. Kramer's estimation technique, the approach is very capable of evaluating the evidence for the single-fault hypotheses.

This would suggest that it may be useful as a refinement method to examine the most likely fault conditions as produced by the diagnostic model processor. In this manner, the single-fault hypothesis can be examined (noting its application) with the ability to return to the full fault set should the results be incorrect or inconclusive. The results shown below can be interpreted with the results generated by the diagnostic model processor indicating that a likely fault choice is a_3. This, however, may not be the only fault occurring as the diagnostic model processor produced high likelihoods of failure for a_1, a_2 and a_4 as well. This reflects the possibility of multiple faults.

The F values are for the single-fault hypotheses as follows:

$$F(a_1) = 1/\sqrt{3} \approx 0.58$$
$$F(a_2) = 1/\sqrt{3} \approx 0.58$$
$$F(a_3) = 1$$
$$F(a_4) = 1/\sqrt{3} \approx 0.58$$
$$F(a_5) = \sqrt{3}/2 \approx 0.87$$

It is obvious that the best single fault hypothesis, besides a_3 which matches the evidence exactly, is a_5. This assumption matches the evidence in all but one constraint, C_4; this is an interesting result considering that it was eliminated as a fault possibility prior to using the estimation technique for evaluating the evidence in Dr. Kramer's example.

Regarding Dr. Kramer's last comment about the calculation of the tolerances of secondary constraints, his correction is statistically accurate. Equation 6 in Petti et al. is used as a more conservative estimate of the secondary tolerance to ensure that an equation is not derived, which would prove to be more severely violated then the parent equations. This is desirable since the tolerances that are used for the primary equations are not determined from a statistical analysis of the constraints.

Thomas Petti
Jim Klein
Department of Chemical Engineering
University of Delaware
Newark, DE 19716

To the Editor:

This letter is in response to the paper by Skogestad, Lundström, and Jacobsen titled "Selecting the Best Distillation Control Configuration" (May 1990, p. 753). In the paper titled "Control Configuration Selection for Distillation Columns" (October 1987, p. 1620), Skogestad and Morari used a spectrum of distillation column cases to analyze control performance since the dynamic behavior and resulting control performance can change significantly with relative volatility, product purity, and feed composition. I like the idea very much.

Distillation columns with a large relative volatility (e.g., $\alpha > 3$) are characterized by very fast dynamic responses since the reflux to feed ratios are usually low. Fortunately, a large relative volatility is accompanied by a relatively large temperature drop across the column which facilitates the use of tray temperature to infer product composition. As a result, the detrimental effect of composition analyzer deadtime (3 to 12 minutes considering industrial sample transport delays) can be negated by controlling tray temperature (which responds quickly) and choosing the setpoints for the tray temperatures from the composition analyzer updates.

Columns with a low relative volatility (e.g., $\alpha < 1.3$) typically use relatively high reflux to feed ratios and are quite slow responding processes. These columns do not have a large enough temperature drop across them to allow the effective use of tray temperatures to infer product composition. But the analyzer deadtime to process time constant ratio is sufficiently low so that the deadtime does not adversely affect control performance.

In light of the above perspective, I have a number of questions about several of the modeling assumptions made in this paper.

I. It is assumed that the molar holdup on each stage to feed rate ratio is a constant 0.5 minute for all cases considered. At a constant feed rate and constant product compositions, as the relative volatility is decreased, the resulting column diameter will increase due to the increase in vapor-liquid traffic. As a result, the molar holdup to feed rate ratio will be larger for low relative volatility columns than for high relative volatility columns. In addition, it is common practice to design a high relative volatility distillation column with a rectifying section with a smaller diameter than the stripping section since the liquid traffic in the rectifying section is typically lower. For example, I have seen a C_2/C_3 splitter for which the molar holdup in the rectifying section was about half of that in the stripping section. The variation in M_i/F between low and high relative volatility columns is significantly larger than the latter example and could amount to an order of magnitude variation in M_i/F for the cases considered in this paper. This difference in M_i/F would cause very significant differences in column dynamics between low and high relative volatility columns unlike the results shown in the paper.

II. The authors assumed that the liquid holdup in the reboiler and the overhead accumulator were also each equal to the molar holdup on a single tray. The authors offer the following reasons for that assumption. "1) If temperatures are used to estimate composition, then the dynamic response *inside* the column matters and this response is usually weakly affected by the reboiler and condenser holdups (Skogestad and Morari, 1988a). 2) The holdups (levels) in the condenser and reboiler may vary with time and it is then safest (from a robustness viewpoint) to use the smallest holdup when designing the controllers (Skogestad and Morari, 1988b)."

With regard to the first reason, when tray temperatures are used to infer product compositions they are usually located a number of trays away from the product end. Also, the residence time of an industrial accumulator can be substantial (as much as an hour). Then, for example, after the reflux rate is changed, the composition on the tray used to infer overhead composition will not reach its final value until the composition of the accumulator has adequately approached its new value. As a result, the molar holdup in the accumulator and reboiler will introduce a major lag into the dynamic response of the process. Simulations based upon a low holdup in the accumulator and reboiler would seem to have a considerably faster response than an industrial column.

For the second point, while a fast responding system is more sensitive to deadtime and nonstationary behavior, distillation columns, which industrially have large reboiler and accumulator holdups, exhibit a highly lagged, overdamped response which is a formidable control problem in their own right. Consider the results for Column D shown in Figure 8 of their paper for a feed composition upset and a feed flow rate change. For the reflux ratio/boilup ratio control, the largest relative error in y was about 2% while the time to return to setpoint was about 15 minutes. Since Column D has a relative volatility of 1.12, one would have to rely upon composition analyzer updates for composition measurements and there would be about a 5-minute analyzer deadtime. In industrial columns of this type, with a normal-sized accumulator, it would take at least 15 minutes before the feedback control was aware that there was a significant deviation from setpoint. The simulated low holdups for the accumulator and little or no deadtime probably contributed to the use of the high controller gains (the L/D, V/B controller gains were each 113 or 30 times that used by the other controllers). With the lag of the accumulator and reboiler, I would expect that the control response for this case would be considerably different since using a high controller gain on a system with a lot of lag can cause oscillatory or even unstable behavior.

III. It was stated that the controller settings developed for setpoint changes were somewhat sluggish for disturbance rejection. This is a common observance with linear controllers. In theory, a SISO PID controller could be applied for disturbance rejection by raising the controller gain in order to keep the process at the point for which it was tuned. In practice, if a PID controller is highly tuned at a point for disturbance rejection, it will sooner or later deviate from setpoint enough to cause poor control performance. As a result, control engineers must "detune" their controllers in order to insure reliable performance.

IV. Four of the cases considered use a relative volatility of 1.5. This corresponds to a ethylene splitter and represents a lower limit on using tray temperatures to infer product compositions. As a result, using tray temperatures to infer product compositions is not as accurate as for the other cases that have higher relative volatilities. It appears to me that case B would pose quite a difficult control problem; and since the feed contains only 10% light material, it would be difficult to use a tray temperature in the bottom part of the column to infer bottoms compositions since there would be a relatively small temperature drop across the stripping section. As a result, one would have to rely upon the composition analyzer for bottom composition feedback. The analyzer deadtime is likely to significantly affect the control performance since the reflux to feed ratio is not particularly high.

V. While I have these questions about the assumptions used in the simulated results, I concur with the author's conclusion that V/B, L/D control should provide the best dual composition control performance. From a McCabe Thiele diagram, one can see that the product purity changes can be affected by changing the slope of the operating line. The slope of the operating line in the stripping section is directly related to V/B while in the rectifying section it is related to L/D. As a result, by maintaining these ratios during upsets, significant decoupling between the ends of the column is provided. For example, if the boilup rate increases, maintaining the vapor/liquid ratio in the rectifying section helps reduce the effect of the boilup rate change on the top of the column. The advantages of this approach can be seen by examining the behavior of LV control to a boilup rate change.

Jim Riggs
Dept. of Chemical Engineering
Texas Tech University
Lubbock, TX 79409

Reply:

In our paper (Skogestad et al., 1990) we demonstrate the effectiveness of the frequency-dependent RGA as a tool for selecting the best distillation column configuration. These results are general, but in the examples we naturally chose specific values for relative volatility, holdups, etc. Although the example columns were chosen to represent a range of columns, it is not possible to cover all combinations of parameters that may be encountered in practice. In particular, we agree that it is unrealistic that the holdups are the same in all example columns. This was done intentionally to keep the presentation clearer.

Nevertheless, we are happy for the comments made by Dr. Riggs because this gives us the opportunity to clarify some issues which we were not able to discuss in our paper. Riggs considers the two cases: a) large relative volatility columns with fast dynamic response; and b) low relative volatility columns with a relatively slow dynamic response. It should be noted, however, that although it is true that low volatility columns generally have a slow *open-loop* response, it does not imply that the response under feedback control has to be slow. This is a myth which is very often encountered among operators. Note that the overall lag for liquid response, θ_L, which is one of the most important dynamic parameters for feedback control of distillation columns, is not generally large for low volatility columns. We agree that the signal to noise ratio makes it simpler to use tray temperatures to infer compositions in case a, but we do not agree that it is not possible to use temperature measurements in case b where the temperature difference between top and bottom is small. In fact, the signal to noise ratio may be improved by adding more temperature measurements as discussed in the following.

I. Holdup on Stages. We agree that the molar holdup to feed rate (M_i/F) usually is larger for columns with a large reflux. There is, however, no general relationship because the holdup also depends on the particular packing or tray chosen. In fact, we have observed similar values of $M_i/F \approx 0.5$ min in real columns where L/F has been as different as 3 and 20. In any case increasing the holdup will simply amount to changing the time scale in the results described in our paper. To obtain the holdup in a practical situation it is probably simplest to observe the time, θ_L, it takes for a change in reflux to affect reboiler level. For packed columns the overall liquid holdup *inside* the column, M_I, may then be estimated as $M_I/L =$

$\theta_L/0.6$ where L is the average reflux in the column (Skogestad and Morari, 1988). For tray columns this will yield the holdup above the weir, and the "dead volume" under the weir must be added to get the total holdup.

II. Holdup in Condenser (Accumulator) and Reboiler. We certainly agree that the holdups used in our example columns are much lower than will be observed in a real column. We believe, however, that the use of negligible holdup in the reboiler and condenser is justified by the two reasons we stated. In addition, it should be noted that the RGA is almost unaffected by these holdups.

The response in the condenser (y_D) and the reboiler (x_B) will have an additional lag approximately equal to the residence time in the condenser and reboiler. This lag is undesirable for control purposes and it will vary depending on the holdup. It is, therefore, better in such cases to measure the compositions inside the column, for example, in the overhead vapor line. Alternatively, temperature measurements may be used. As mentioned in the paper they have a dynamic response which is usually weakly affected by the condenser and reboiler holdup.

1. It is not correct, as stated by Riggs, that the holdup in the accumulator and reboiler will introduce a major lag into the dynamic response inside the column. The reason is that the compositions inside the column will start changing immediately if, for example, the liquid flow changes.

2. The seemingly high controller gains for the $(L/D)(V/B)$ configuration are mainly a result of input scaling. For example, if the essentially equivalent $[L/(L+D)][V/(V+B)]$ configuration was used instead, the controller gains for column D would be reduced by a factor 413 (top) and 1,110 (bottom).

III. Disturbance Rejection and Set Points. The fact that the controller settings developed for set point changes were somewhat sluggish for disturbance rejection has nothing to do with the use of linear controllers. The same results would apply to nonlinear controllers. The reason is simply that the disturbances are large and to reject them we require fast response by the controller. For set-point changes these tunings give a response that is faster than what we actually want. In such cases, the feedback controller should be designed for disturbance rejection and a prefilter should be added to the set-points to dampen the response (two-degree-of-freedom controller). Furthermore, for distillation columns the effect of nonlinearity caused by deviations from the set-point may be almost eliminated by using logarithmic compositions as inputs to the controller (Skogestad and Morari, 1988).

IV. Use of Temperatures to Infer Compositions. If only one temperature is used to infer composition, it may be correct that a relative volatility of 1.5 represents a lower limit. We propose, however, that all available measurements along the column should be used when estimating compositions. In this case, it is possible to get a good estimate of the compositions even when the temperature difference across the column is relatively small.

V. Intuitive Arguments for Ratio Configurations. The McCabe-Thiele arguments presented by Riggs are intuitively appealing, but it is incorrect to believe that one can deduct control properties using these arguments. First, the arguments apply to steady state only; and as the DB configuration shows, these may be totally misleading for evaluating control performance. Second, even at steady state, the argument may be misleading (Skogestad and Morari, 1987). For example, it may be shown that the control properties of the LV and $(L/V)V$ configurations are essentially the same, even though the latter has a ratio as one of the inputs. As for Riggs' final comments regarding a boilup rate disturbance, this is exactly the information contained in Table 3 of our paper.

Literature cited

Skogestad, S., and M. Morari, "Control Configuration Selection for Distillation Columns," *AIChE J.* **33**, 1620 (Oct. 1987).

———, "Understanding the Dynamic Behavior of Distillation Columns," *Ind. Eng. Chem. Res.*, **27**, 1848 (Oct., 1988).

Skogestad, S., P. Lundström, and E. W. Jacobsen, "Selecting the Best Distillation Control Configuration," *AIChE J.*, **36**, 753 (May, 1990).

Sigurd Skogestad
Petter Lundström
Elling W. Jacobsen
Dept. of Chemical Engineering
Norwegian Institute of Technology
N-7034 Trondheim, Norway

To the Editor:

In the R&D note titled "Use of High-Gradient Magnetic Fields for the Capture Ferritin" (December 1989, p. 2058), Scott described an experimental investigation of the capture of ferritin particles by high-gradient magnetic separation (HGMS). It was concluded that the efficiency of capture was a function of molecular charge and the standard trajectory theory of the process is inadequate.

A perusal of the text quickly reveals that the author is not familiar with recent and not-so-recent developments in HGMS. This is regrettable, since this interesting problem deserves an expert approach.

It was shown about ten years ago (Collan et al., 1979; Svoboda, 1981) that surface forces may have a substantial effect on the performance of a magnetic separator for particles as large as 20 μm. Further experimental and theoretical research into the role of surface interactions in magnetic separation and flocculation (Svoboda, 1982; Svoboda and Zofka, 1983; Parker et al., 1982; Parsonage, 1984; Svoboda et al., 1986; Svoboda et al., 1987) showed that optimum selectivity and efficiency of magnetic separation of colloidal particles are achieved by control of the surface properties of various components of a suspension. The results imply that the best performance of magnetic-separation process can be expected at the pH value at which the surface potentials of the components to be separated have the same sign. However, this pH value should not be too far from the point-of-zero charge of the component to be recovered magnetically.

Moreover, it has been shown (Svoboda and Corrans, 1985) that the desorption of quasicolloidal particles from the matrix of a magnetic separator is strongly dependent on the pH value of the flushing fluid. A total wash-off of particles cannot be achieved without affecting the surface interactions between the adsorbed particles and the matrix.

It is obvious that such an important

effect of surface forces on the efficiency of magnetic separation of quasicolloidal particles will be considerably more pronounced in the subcolloidal region. It is likely that for 125 Å dia. ferritin particles (Scott, 1989), the surface forces will dominate the balance of forces in a magnetic separator. It has actually been shown (Svoboda, 1987; Svoboda and Ross, 1989) that van der Waals and electric double layer interactions are dominant for particles of diameter smaller than 200 Å and can induce the capture of even diamagnetic particles. It is therefore obvious, from first principles, that the application of the so-called "standard trajectory theory" to magnetic separation of subcolloidal particles is not permissible.

In view of its considerable practical and theoretical importance, the project described in Scott's paper would deserve an enlightened treatment

Literature cited

Collan, H. K., J. Jantunen, M. Kokkala, and A. Ritvos, "Inversion of the Breakthrough Curve of a High-Gradient Magnetic Filter—Theory and Experiments," *Proc. Ind. Appl. Mag. Sep.*, Y. A. Liu, ed., Rindge, NH, 175 (1979).

Parker, M. R., R. P. A. R. van Kleef, H. W. Myron, and P. Wyder, "Particle Aggregation in Colloids in High Magnetic Fields," *J. Mater.*, **27**, 250 (1982).

Parsonage, J., "Selective Magnetic Coating for Mineral Separation," *Trans. Inst. Min. Metall.*, C37 (1984).

Svoboda, J., "A Theoretical Approach to the Magnetic Flocculation of Weakly Magnetic Minerals," *Int. J. Min. Proc.*, **8**, 377 (1981).

———, "The Influence of Surface Forces on Magnetic Separation," *IEEE Trans Mag.*, **MAG-18**, 862 (1982).

Svoboda, J., and J. Zofka, "Magnetic Flocculation in Secondary Minimum," *J. Col. Interf. Sci.*, **94**, 37 (1983).

Svoboda, J., and I. J. Corrans, "The Removal of Particles from the Matrix of a High-Gradient Magnetic Separator," *IEEE Trans. Mag.*, **MAG-21**, 53 (1985).

Svoboda, J., I. J. Corrans, and M. Spitze, "The Effect of pH on the Recovery of Uranium and Gold by HGMS," *Int. J. Min. Proc.*, **17**, 83 (1986).

Svoboda, J., M. Lazer, and W. A. M. te Riele, "Selective Magnetic Separation of Uranium and Gold," *IEEE Trans. Mag.*, **MAG-23**, 283 (1987).

Svoboda, J., *Magnetic Methods for the Treatment of Minerals*, Elsevier, Amsterdam (1987).

Svoboda, J., and V. E. Ross, "Particle Capture in the Matrix of a Magnetic Separator," *Int. J. Min. Proc.*,. **27**, 75 (1989)

Jan Svoboda
Magnetic Technology Consultants (Pty) Ltd
Vereeniging 1930, South Africa

Reply:

The results described in this preliminary study were intended to spark interest in the separations community for extension of HGMS techniques for possible use in separations of charged molecules. For brevity, only facts pertaining to the reported experimental results were included in the text. As this area is a nonstandard one in chemical engineering circles, the closing remarks were directed toward the (obvious) conclusions in such a manner so as to relate the significance of the findings in terms of the most widely understood approach (the trajectory analysis). There was no intent to present a fundamentally-based explanation of the observed phenomena; hence, there was no ill intent toward researchers who may have addressed related issues.

Since no attempt was made to address theoretical issues or specify our ongoing approach to work in this area, I find it unfortunate that comments should be made regarding these items, particularly when comments are rendered by persons who have not contacted us regarding our work. The current experimental approach involves a split-coil superconducting magnet which allows both optical and spectroscopic access to suspensions in various magnetic field configurations. This arrangement allows a much more rigorous inspection of particle behavior than in the reported system. The theoretical approach being taken is fashioned after that of Karrila and Kim (1989), in which transient interactions involved in fluid-particle motion and charge and magnetic field effects are all taken into account. This technique also has the capability of addressing multiparticle systems where the effects of multiparticle interactions have a significant bearing on behavior of the suspension.

The approaches touted by the author of the above letter involve highly idealized treatments which accomplish little more than a qualitative description of expected behavior. While useful for understanding trends in some experimental data, this hardly represents a fundamental approach to the problem.

Literature cited

Karrila, S. J., and S. Kim, "Integral Equations of the Second Kind for Stokes Flow: Direct Solution for Physical Variables and Removal of Inherent Accuracy Limitations," *Chem. Eng. Comm.*, **82**, 123 (1989).

Timothy C. Scott
Chemical Technology Division
Oak Ridge National Laboratory
Oak Ridge, TN 37831