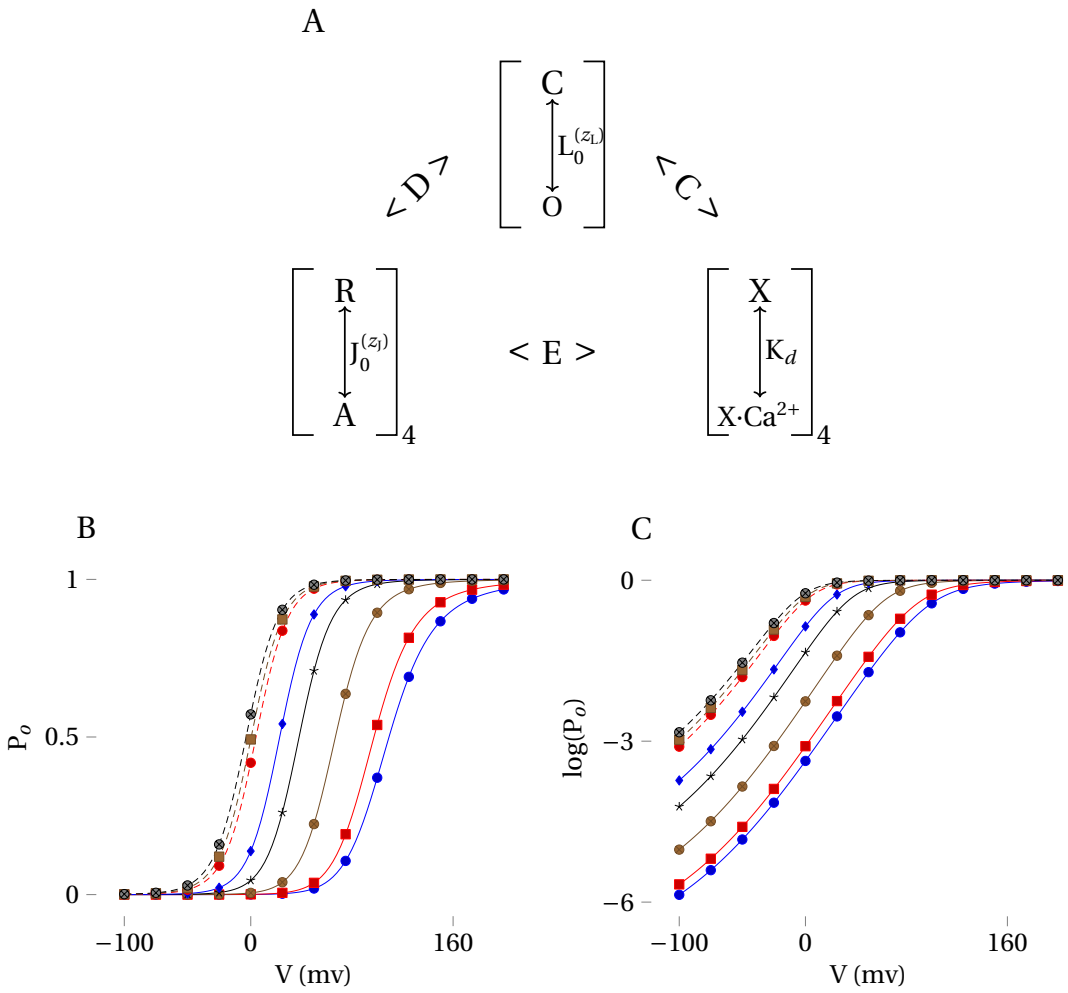
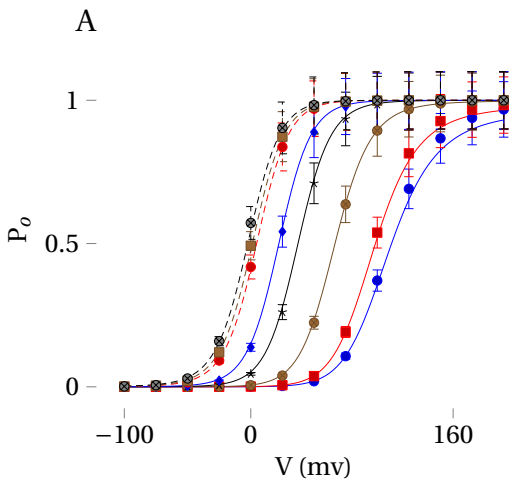
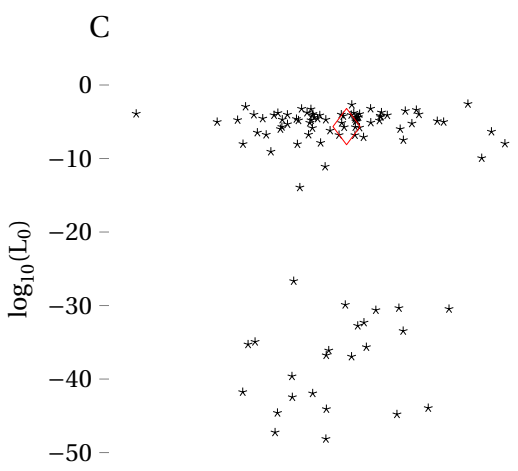


Figure 1: Synthetic steady-state data. (A) Schematic of the general allosteric gating mechanism used to generate synthetic data. The steady state properties of this model are fully described by eight parameters, three of which define the allosteric interactions (C, D, E) and the remaining five define the equilibrium constants (J,K,L) via J = $f(J_0, z_J)$, K = $f(K_D)$, L = $f(L_0,\ z_L)$. (B, C) $P_o - V$ and $\log(P_o) - V$ relationships generated from the scheme in (A) for different Ca^{2+} (in μm : 0 (●); 0.7 (■); 4 (●); 12 (⋆); 22 (♦); 55 (●); 70 (■); 95 (⊗)) using previously published best-fit parameters . Each curve contains 26 data points, sampled from the underlying binding curve (solid lines).



B

θ_i	Base	Fit
L_0	2.2×10^{-6}	7.7×10^{-37}
z_L	0.42	8.8×10^{-33}
J_0	0.10	0.066
z_J	0.58	0.59
K_d	3.9×10^{-5}	2.7×10^{-5}
C	6.16	1.7×10^{8}
D	30.4	2.4×10^{9}
E	2.0	4.5×10^{-8}



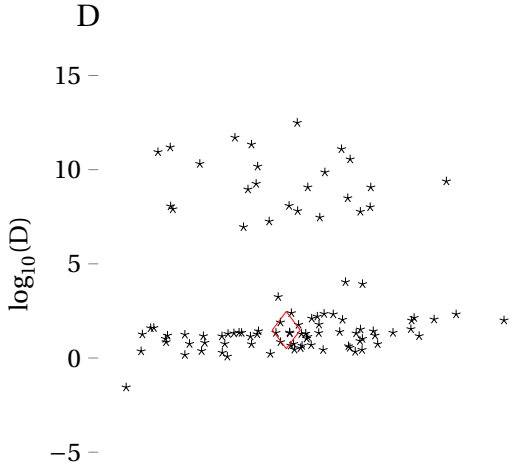


Figure 2: An illustration of non-identifiability. (A). Data generated with 'base' parameters, well-fit with 'Fit' parameters. Error bars are 10% from data value. (B). Parameters which generated the data ('base') and the solid lines ('fit') shown in (A). (C,D). Log of fitted L_0, D (respectively) values to 100 noisy synthetic P_o datasets generated from 'Base' parameter values. Values span many orders of magnitude. x-axis spread is for ease of visualization, red diamond surrounds the true 'base' parameter.

A

Parameter (θ_i)	L_0	z_L	J_0	z_J	K_d	C	D	E
P_o Relative error (%)	5×10^{11}	4.0×10^4	7.4×10^3	71.80	304.31	2.8×10^3	1.8×10^6	1.7×10^4
$\log(P_o)$ Relative error (%)	7.9×10^3	612.18	262.10	9.50	13.14	32.88	1.2×10^3	22.19

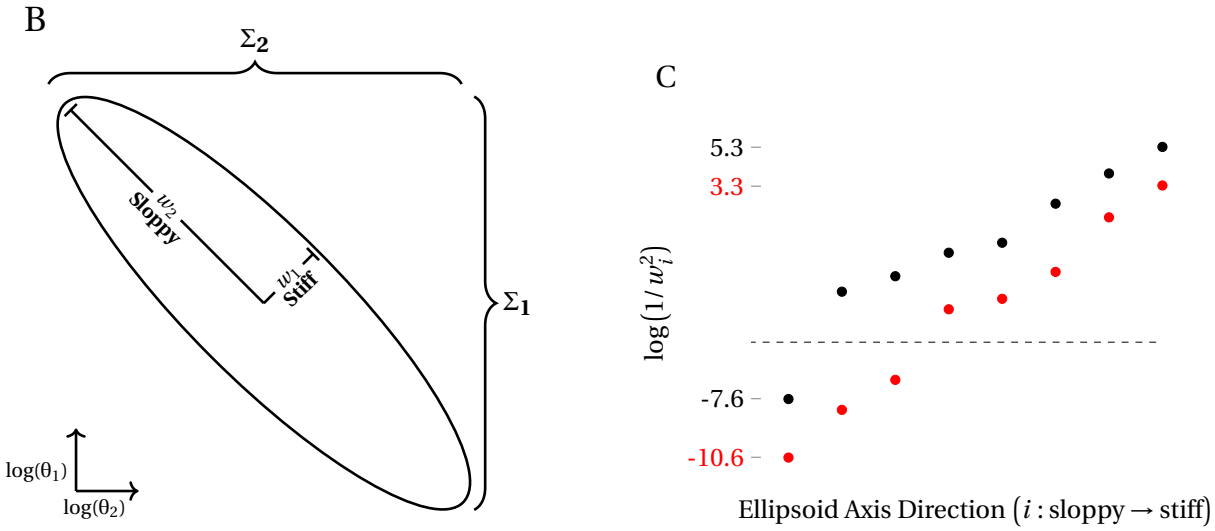


Figure 3: The BK model is sloppy. (A) Lower bounds on parameter error (95% confidence interval) for each of the $P_o, \log(P_o)$ assays. The P_o assay exhibits much worse identifiability than the $\log(P_o)$ assay. (B). Ellipsoid of constant cost for a toy two-parameter model. The parameters $\theta_{1,2}$ are constrained in the stiff direction, but have large error regions $\Sigma_{1,2}$ due to the presence of a large sloppy direction. (C) Calculated $\log(1/\text{width}_i{}^2)$ values for the P_o assay (red) and $\log(P_o)$ assay (black). Both exhibit a linear trend, the signature of a sloppy model.

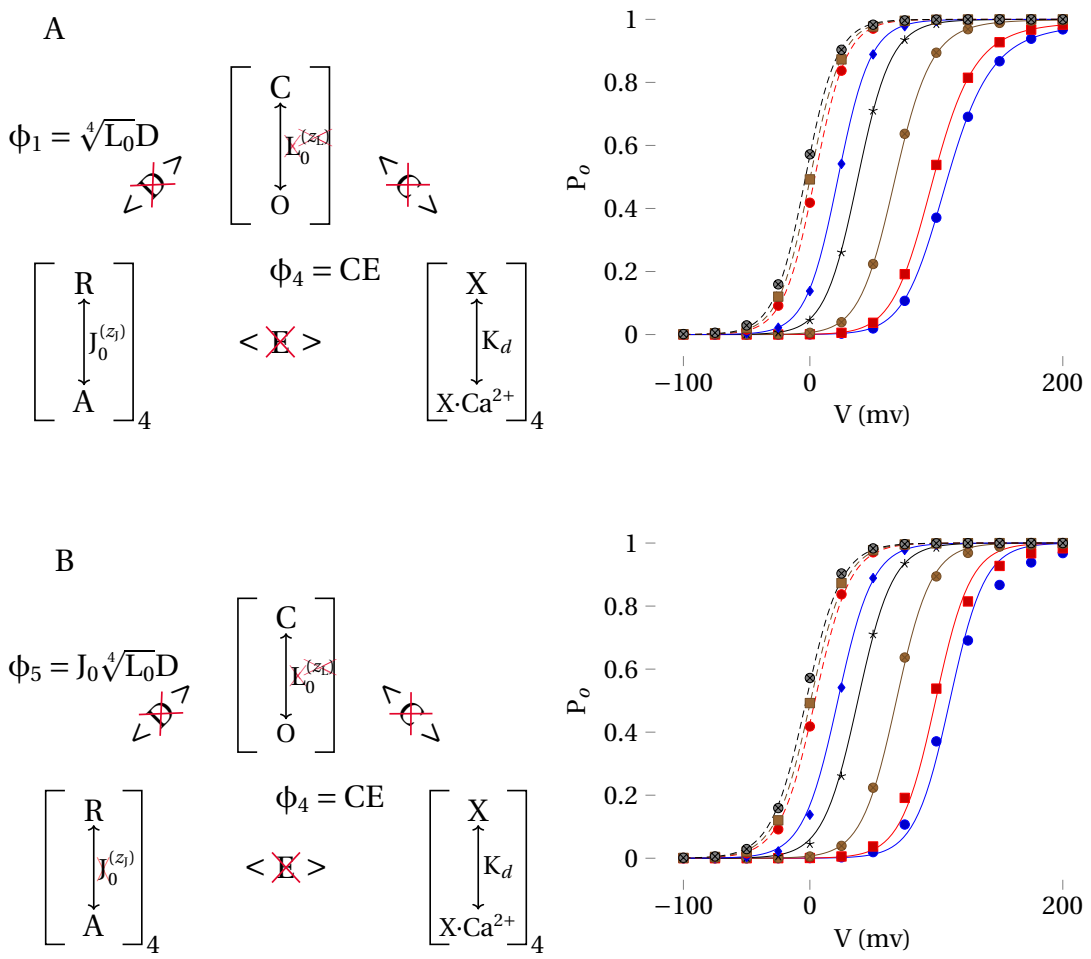


Figure 4: Results of model reduction for P_o. (A, left) Schematic of the model admitted by the third reduction; five parameters have been eliminated from the original model (red X) and two new 'emergent' parameters have been added ($\phi_{1,2}$) for a total reduction of three parameters. (A, right). This model fits the data (solid lines) extremely well. (B, left). Schematic of model admitted by the fourth reduction. This model does not fit the data well at low Ca^{2+} (B, right). Synthetic data labeled as previously.

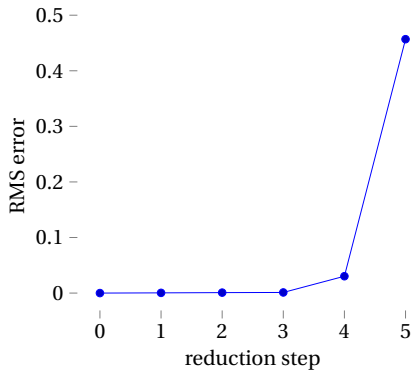


Figure 4 - supplement figure 1: RMS cost for each reduced model. Note a small uptick in error for the fourth reduction, consistent with the poor fit to low [Ca^{2+}] in Figure 4B. Subsequent reductions do not fit the data well.

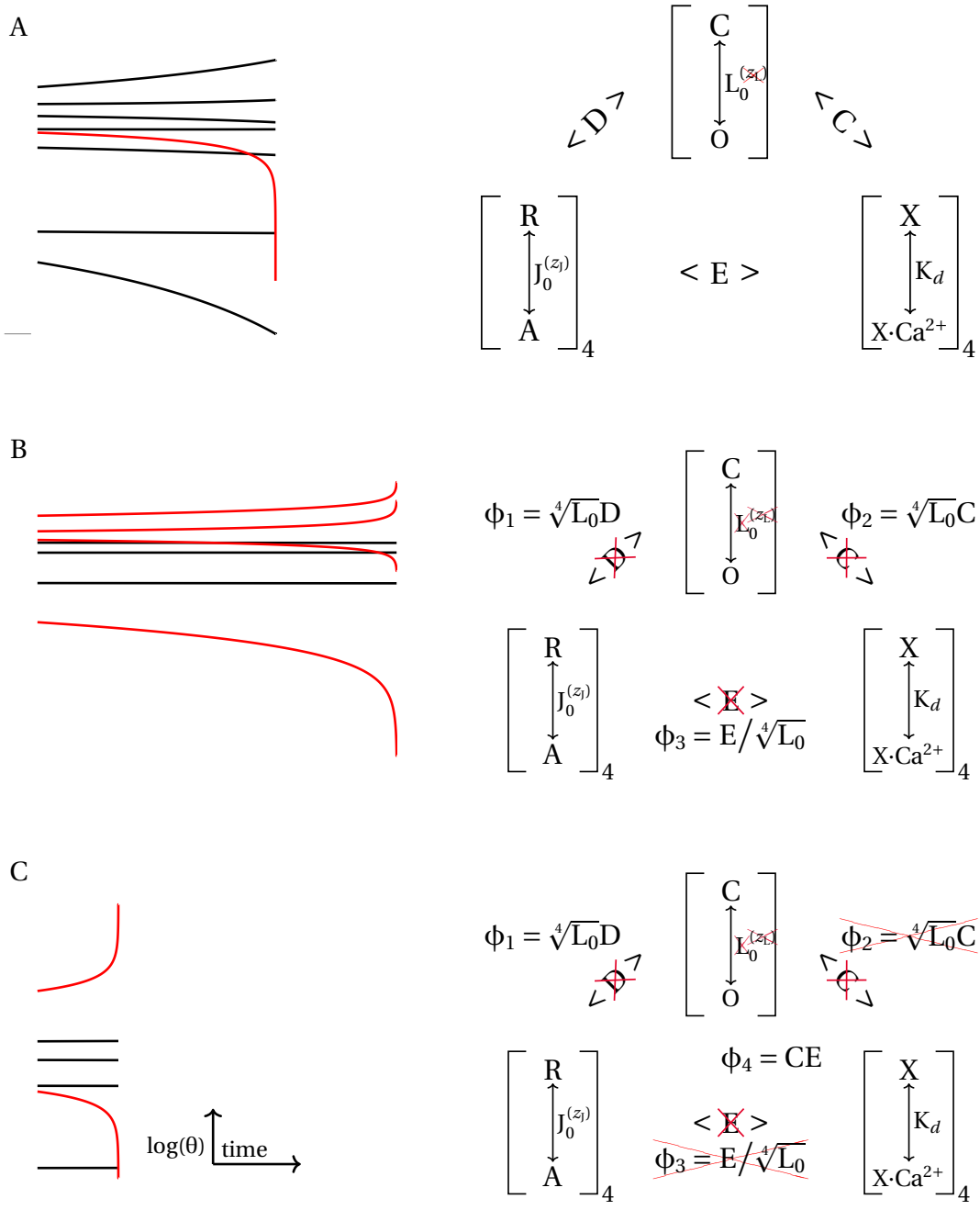
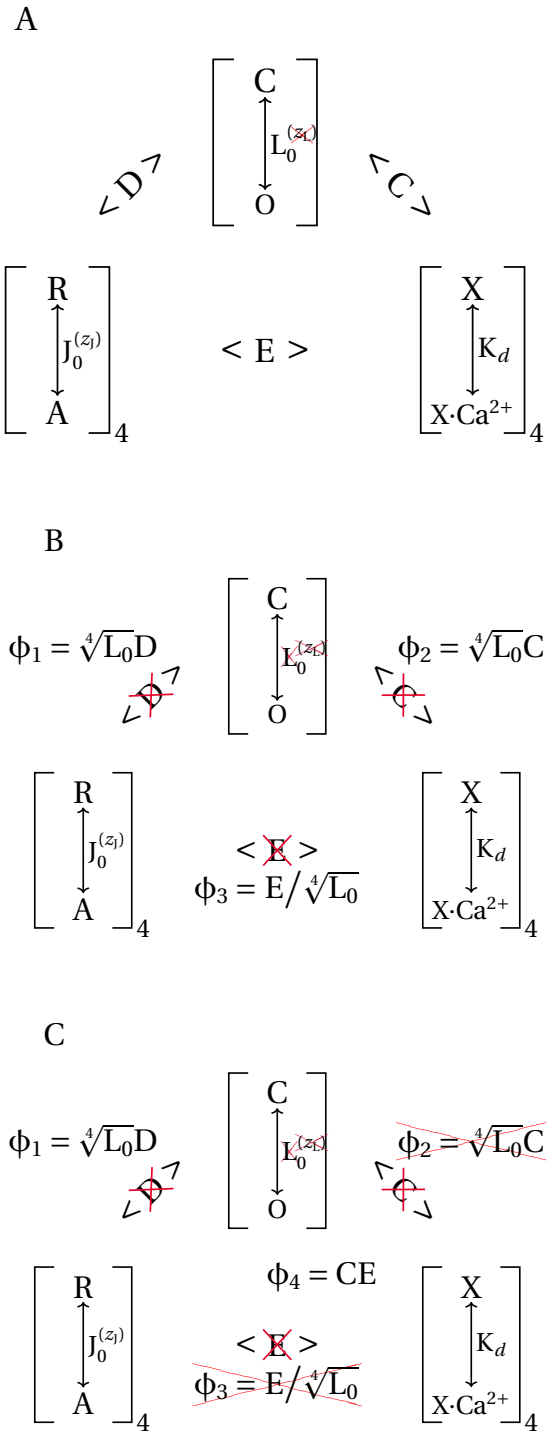


Figure 5: Intermediate MBAM steps, P_o assay. The figure should be read left to right, top to bottom. The left column displays the parameter values for a given model as MBAM progresses. The reduced model created upon completion of the parameter search is displayed on the right. (A). MBAM run for the full, original model. There are eight lines, corresponding to eight parameters. One of the parameters goes to zero; this is z_L, and it is eliminated, giving our first reduced model, at right. (B). In the second iteration, four parameters are observed to diverge: L_0, D, E, C. These parameters are eliminated, and three new, emergent parameters are created ($\phi_{1,2,3}$), yielding a net reduction of one parameter. (C). Two parameters are observed to diverge: ϕ_2, ϕ_3. Note that there are only six lines, corresponding to the six remaining parameters. The resulting model (right) has five parameters, and fits the data well (Figure 4).



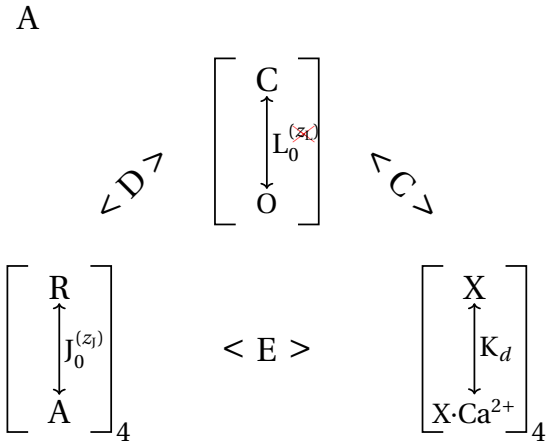
θ_i	Relative error (%)
L_0	1.5×10^{22}
J_0	92.67
z_J	59.82
K_d	128.44
C	1.82×10^{6}
D	1.85×10^{7}
E	7.58×10^{6}

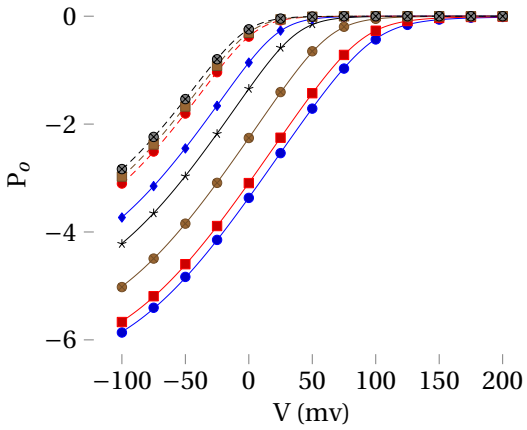
θ_i	Relative error (%)
J_0	82.35
z_J	36.15
K_d	51.03
$\sqrt[4]{L_0}C$	1.23×10^{9}
$\sqrt[4]{L_0}D$	96.86
$\frac{E}{\sqrt[4]{L_0}}$	1.31×10^{9}

θ_i	Relative error (%)
J_0	84.52
z_J	17.99
K_d	47.89
CE	44.78
$\sqrt[4]{L_0}D$	68.67

Figure 6: Model reduction results in identifiable parameters. Reduced models for the P_o assay are presented in the left column, and the error in their parameters (95% confidence interval) are presented at right. (A),(B),(C) correspond to model-error pairs after one, two, and three reduction steps, respectively. The five parameter model produced by three model reductions (C, left) has identifiable parameters (within one order of magnitude for all parameters, right).

A





B

Parameter (θ_i)	L_0	J_0	z_J	K_d	C	D	E
Relative error (%)	68	43	5.6	11	16	52	23

Figure 7: Single reduction identifiability for log(P_o). A model resulting from one reduction (A, left) fits synthetic data (A, right, data legend as in Fig 1) very well (solid lines). This model has identifiable parameters (B).

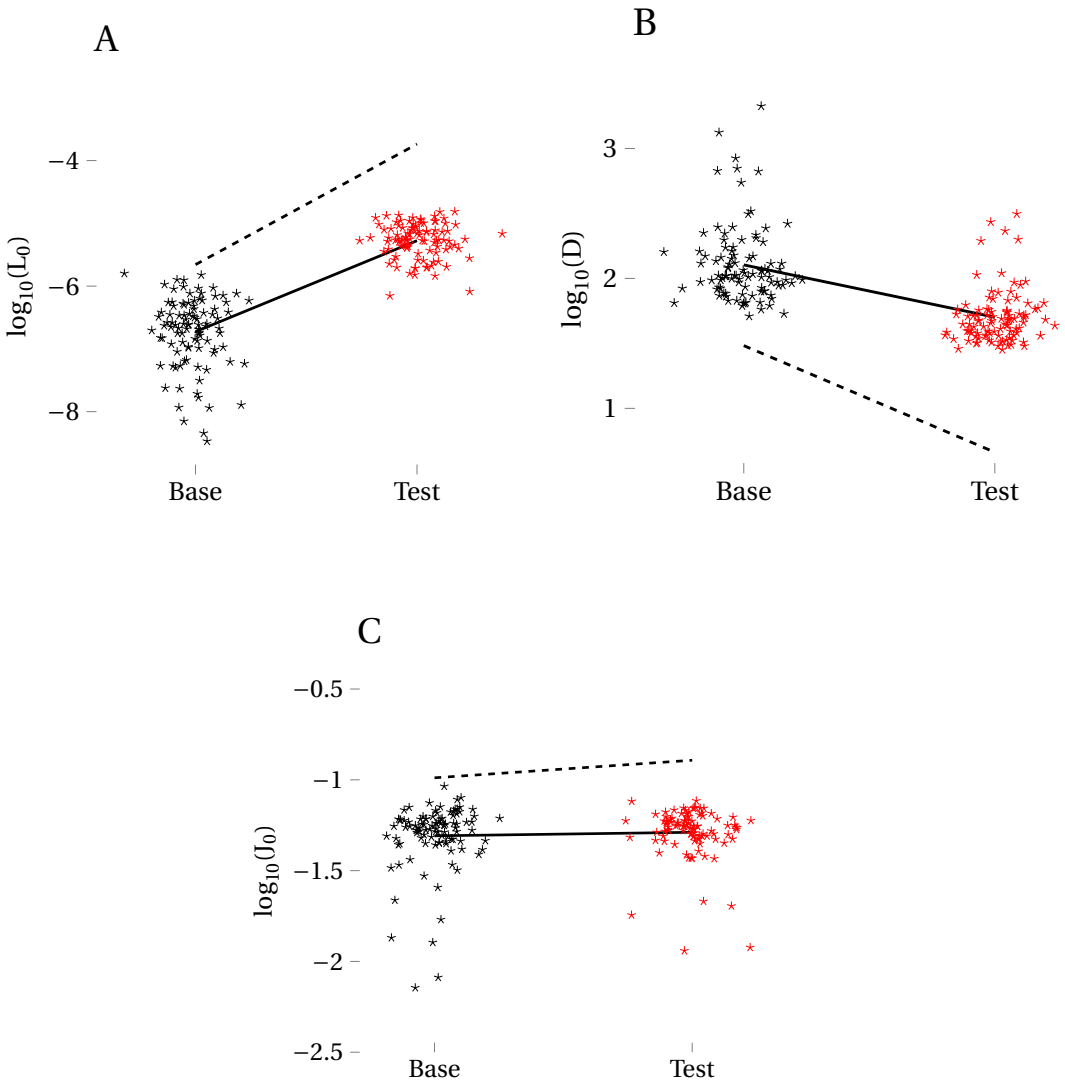


Figure 8: Inferred parameters from the reduced log(P_o) model. Values inferred by fitting the once-reduced log(P_o) model to 100 noisy synthetic measurements generated from base (black) and test (red) parameter sets. (A),(B),(C) correspond to inferred values of J_0, D, J_0, respectively. Solid lines connect the means of the inferred base and test values; dotted lines connect the true generating base and test values.

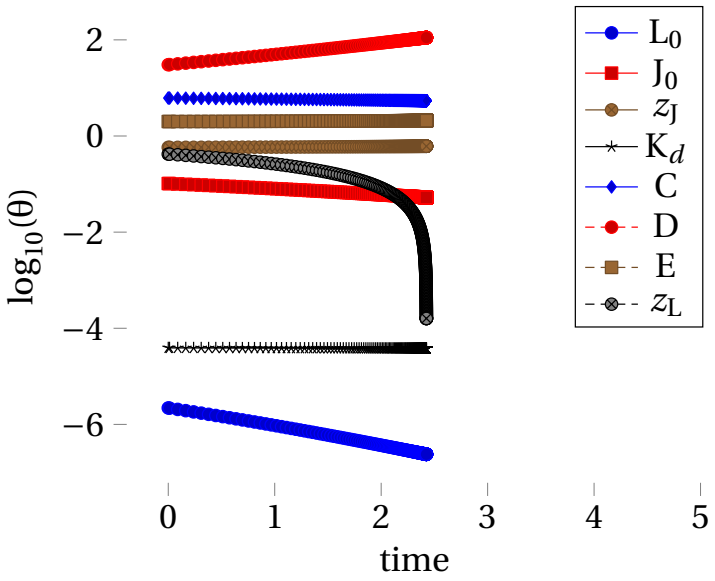


Figure 8 - supplement figure 1: Parameter compensations explain discrepancy between true and inferred values. This figure shows parameter values during the first parameter reduction, which results in the model with one reduced parameter (z_L eliminated). Note that as z_L goes to zero, L_0 and J_0 must decrease, and D must increase, to compensate. We therefore expect that when the once-reduced model is fit to data generated from the full model, L_0, J_0 will be smaller than the true parameters, and D will be larger than the true parameters. This is what we observe in Figure 8.

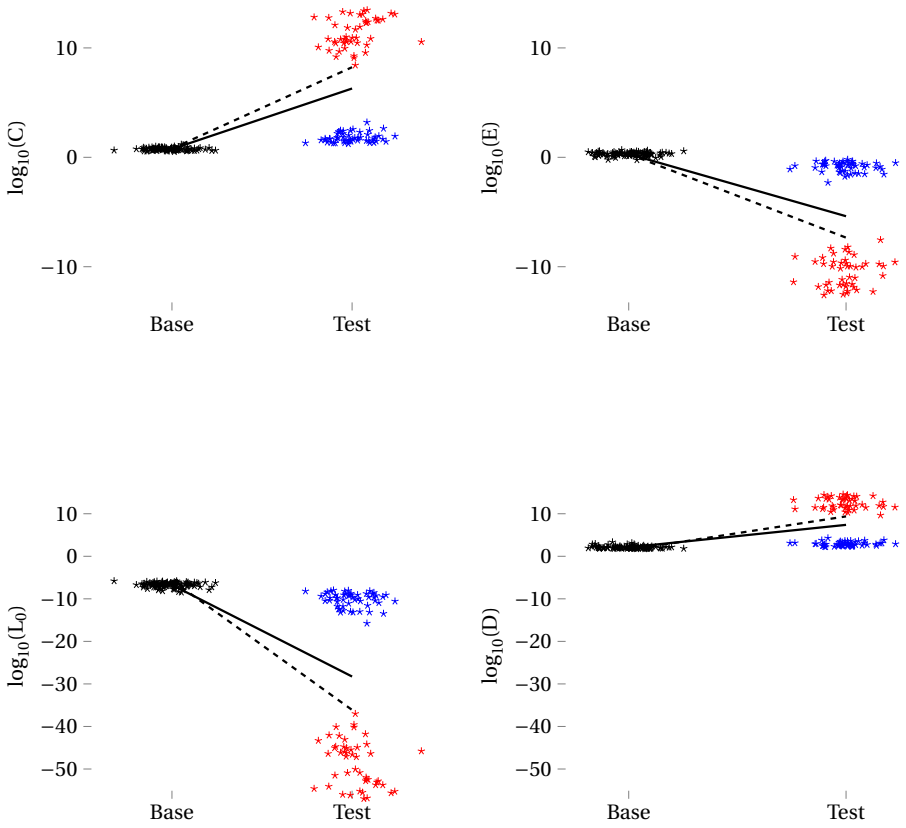


Figure 8 - supplement figure 2: Inferred parameters from the reduced log(P_o) model. Values inferred by fitting the once-reduced log(P_o) model to 100 noisy synthetic measurements generated from base (black) and test (red, blue) parameter sets.. Test parameters are different from those in Figure 8. Here, the inferred values of our compensatory parameters fall into two modes. This is because the model is degenerate at the true 'best fit' parameter values that the noisy synthetic data were generated from. Each set of inferred parameters was labeled either red or blue. The two clouds therefore show that the parameter compensation is consistent for all fits. Each cloud is discernible from the base parameters, and the mean of the values is discernible from the base parameters.

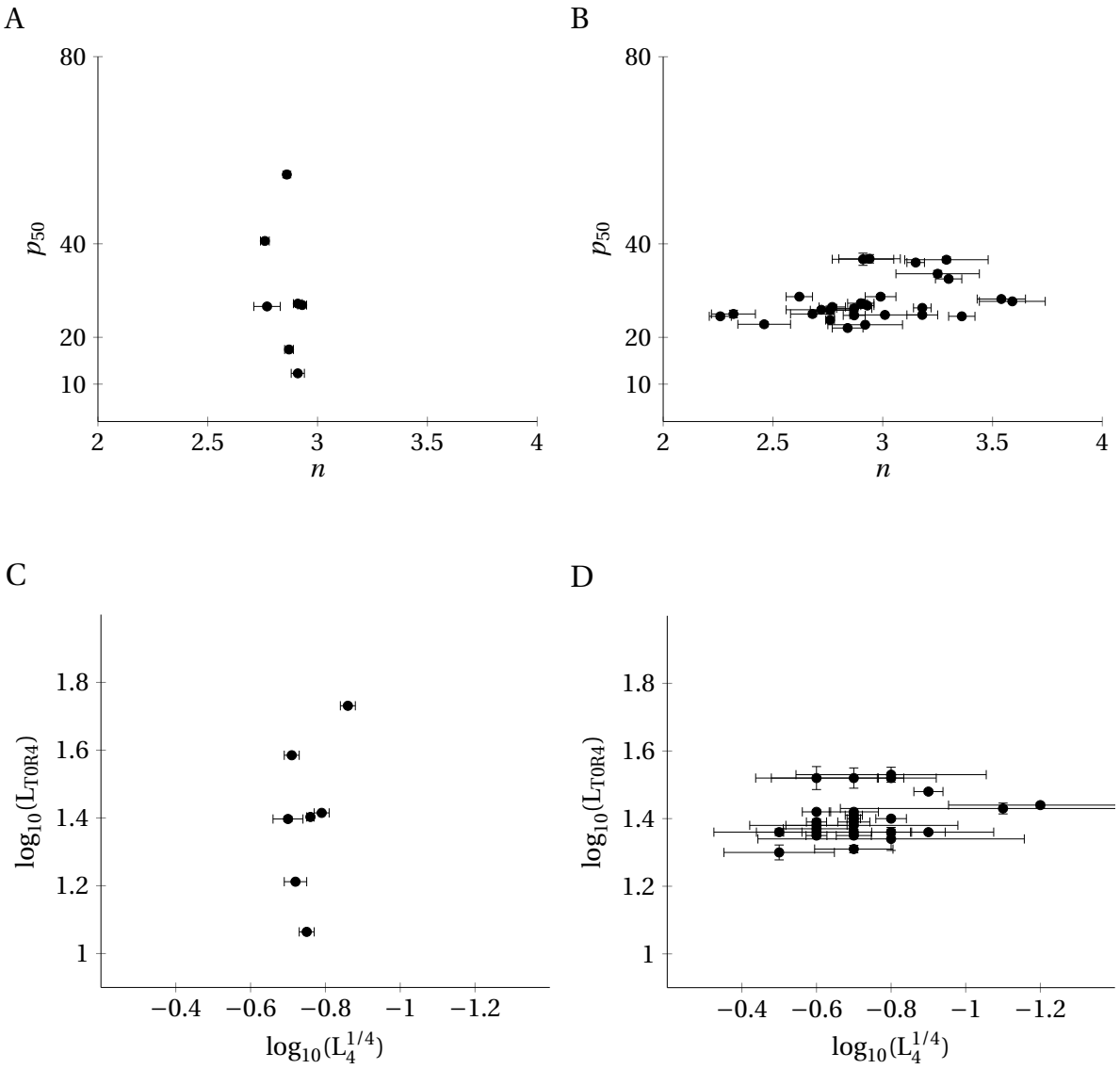


Figure 9: Analysis of hemoglobin oxygen saturation curves. These figures were generated with data published by Milo et al. The Figures aim to reproduce Milo et al's figure 2a,b (our Figure 8A,B) and Milo et al's figure 4a,b (our Figure 8C,D). Each point in Figures 8A,C corresponds to human hemoglobin at a different pH. Each point in Figures 8B,D corresponds to hemoglobin from a different mammal, at the same physiological condition. We refer the reader to the appropriate figures in Milo et al. for full descriptions of pH conditions and mammals used.