Reducibility, identifiability, and evolvability in allosteric macromolecules

November 24, 2016

Abstract

Cell signaling relies crucially on the ability of macromolecules to transduce stimulus information into conformational changes. These conformational changes are often allosteric: one region of the molecule undergoes structural rearrangement in response to stimulus applied at a different region of the same molecule. Here, we address the issue of sensitivity of allosteric macromolecules to their underlying biophysical parameters. A canonical Monod-Wyman-Changeux (MWC) model of the *mSlo* large-conductance Ca^{2+}-activated K^+ (BK) ion channel is observed to have 'non-identifiable' parameters with respect to two common functional assays: neither experimentation provides sufficient constraining power to uniquely estimate model parameter values. We address this non-identifiability by constructing a 'reduced' allosteric BK model for each of the two assays, using the recently developed Model Boundary Approximation Method (MBAM). Each reduced model has fewer total parameters than the original MWC model, but fits its data equally well. Crucially, the parameters of these reduced models are identifiable, and explicitly expressed as emergent combinations of the original MWC parameters. The reduced models thereby identify which coordinated changes in parameter values leave the model output unchanged. We predict that these coordinated changes are evolutionarily relevant 'neutral spaces,' which the protein can use to explore new functions without sacrificing its current behavior. We argue that the biophysical parameters of allosteric macromolecules should be non-identifiable in order to facilitate their evolvability, and discuss how this idea can be interrogated experimentally.

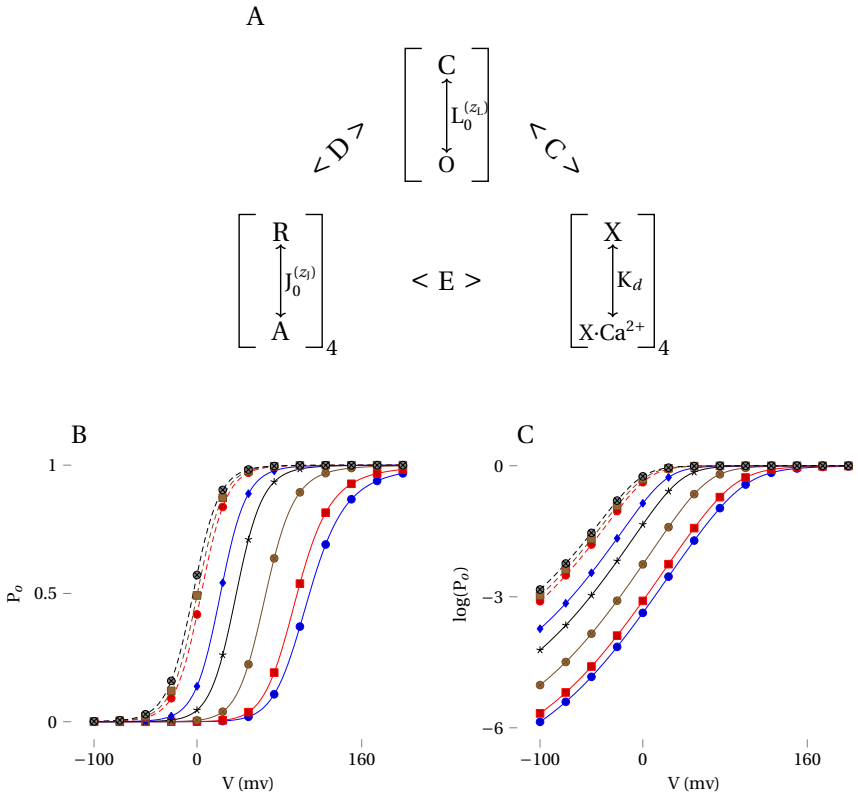
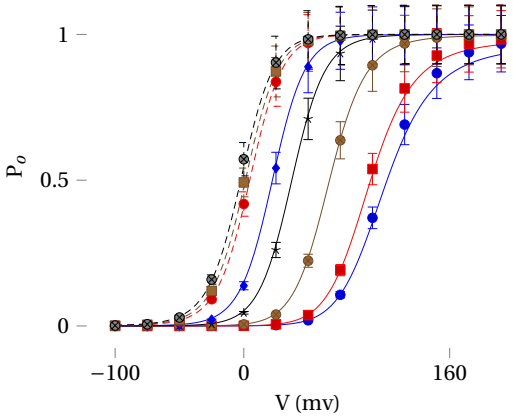


Figure 1: Synthetic steady-state data. (A) Schematic of the general allosteric gating mechanism used to generate synthetic data. The steady state properties of this model are fully described by eight parameters, three of which define the allosteric interactions (C, D, E) and the remaining five define 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, linearly connected for ease of visualization.

A



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: test

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

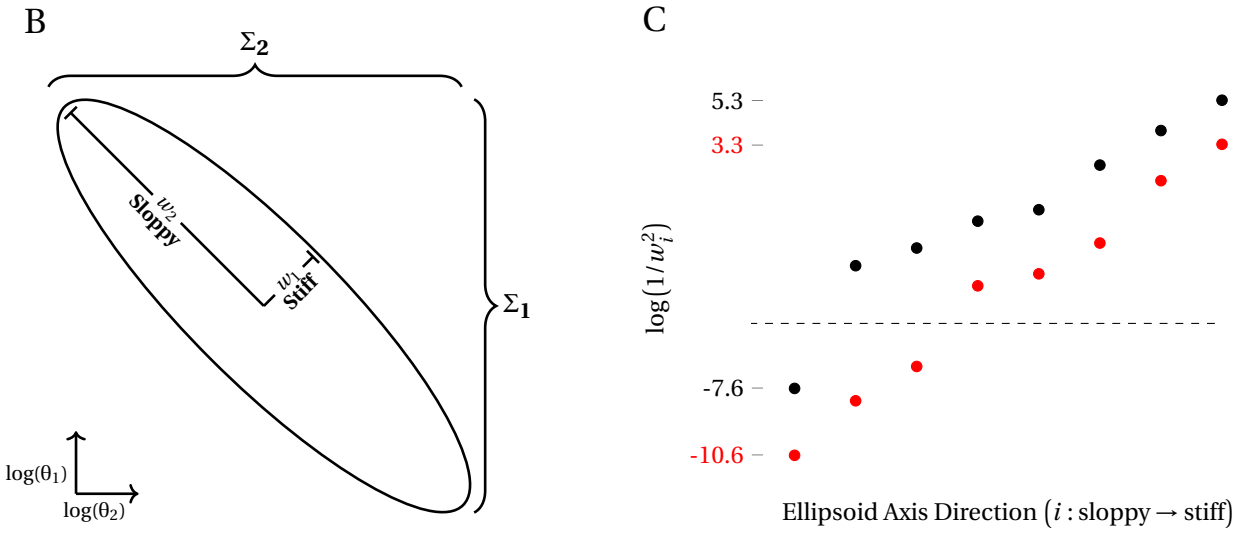


Figure 3: The BK model is sloppy. (A) Ellipsoid of constant cost for a toy two-parameter model. The parameters $\theta_{1,2}$ are constrained in the stiff direction, but have large error regions $\Sigma_{1,2}$ due to the presence of a large sloppy direction. (B) Calculated $\log(1/\text{width}^2)$ values for the P_o scheme (red) and $\log(P_o)$ scheme (black). Both exhibit a roughly linear trend, which is the hallmark of a sloppy model. (C) Parameter error (95% confidence interval) for each of the schemes. The P_0 scheme has most parameters with great errors (>10%), whereas the parameters of the $\log(P_o)$ scheme are all identifiable except L_0.

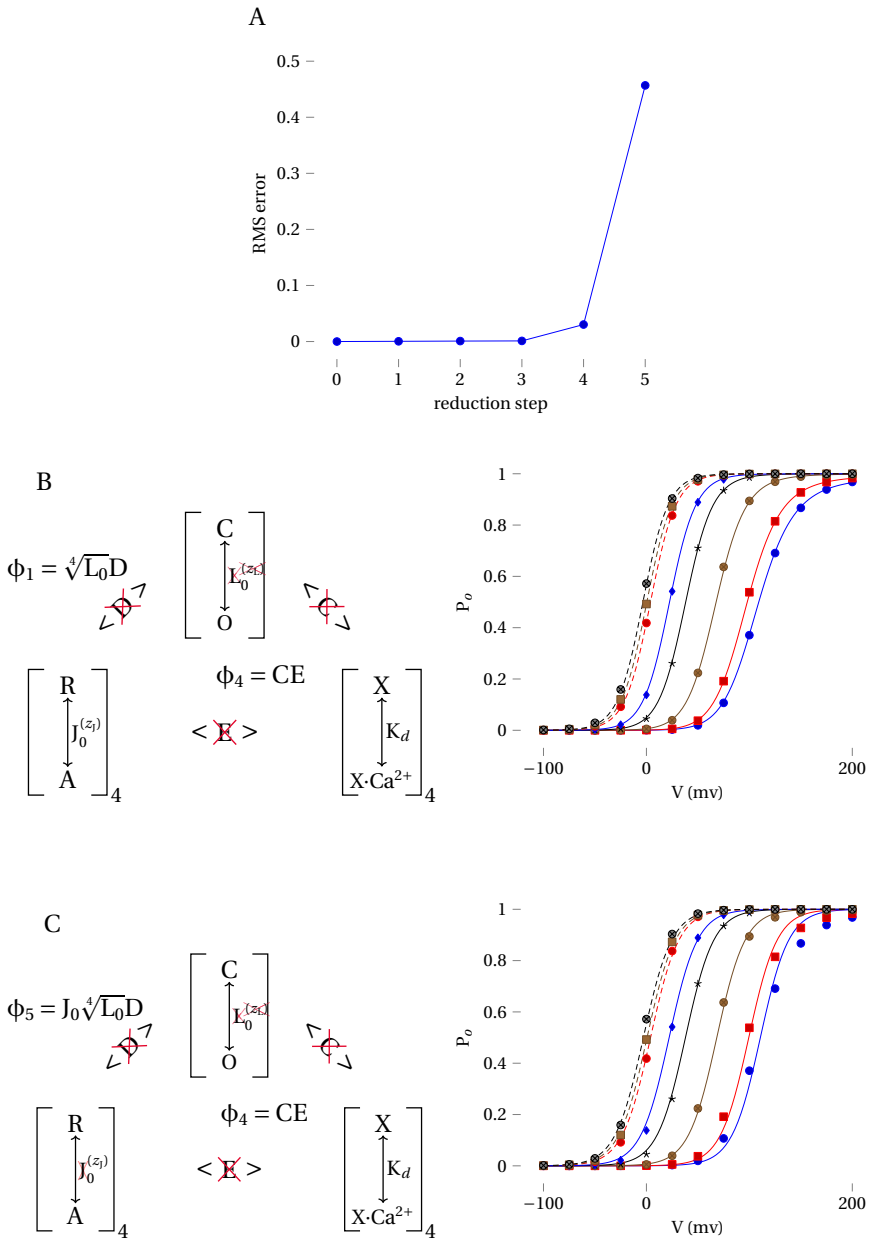


Figure 4: Overview of model reduction for P_o. (A) Error for five reduction steps. Each reduction results in a new model with one fewer parameter. (B, left) Schematic of the model admitted by the third reduction (left); five parameters have been eliminated from the original model (red X) and two new 'emergent' parameters have been added ($\phi_{1,2}$) for a total reduction of three parameters. (B, right) this model fits the data (black lines) extremely well. (C, left) Schematic of model admitted by the fourth reduction. This model does not fit the data well at low Ca^{2+} (C, right), consistent with the increase in computed error (A).

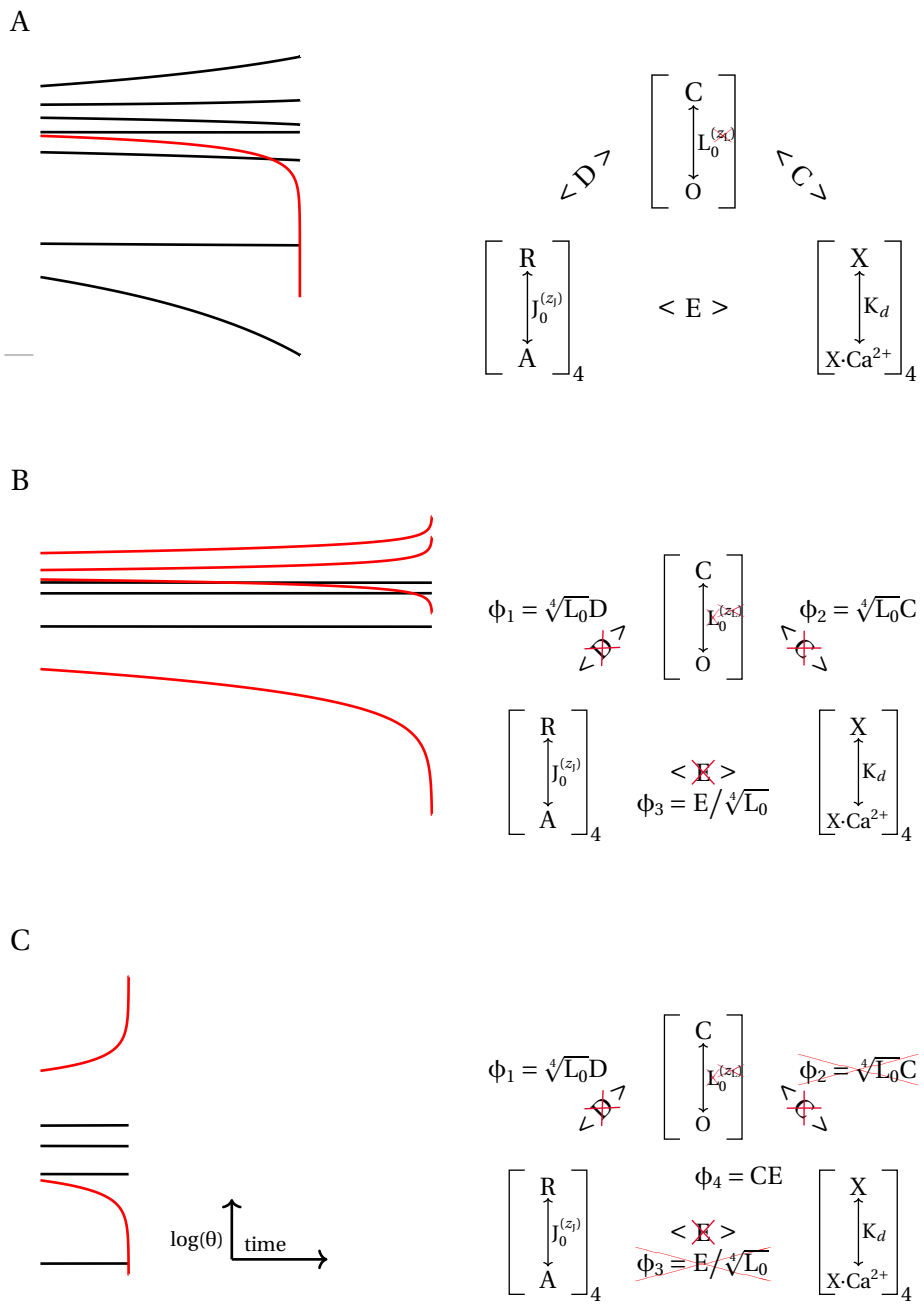


Figure 5: Intermediate MBAM steps, P_o assay. The figure should be read left to right, top to bottom. The left column displays the parameter values for a given model as MBAM progresses. The reduced model created by an MBAM iteration is displayed on the right. (A). MBAM run for the full, original model. Note that 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 (Fig 3).

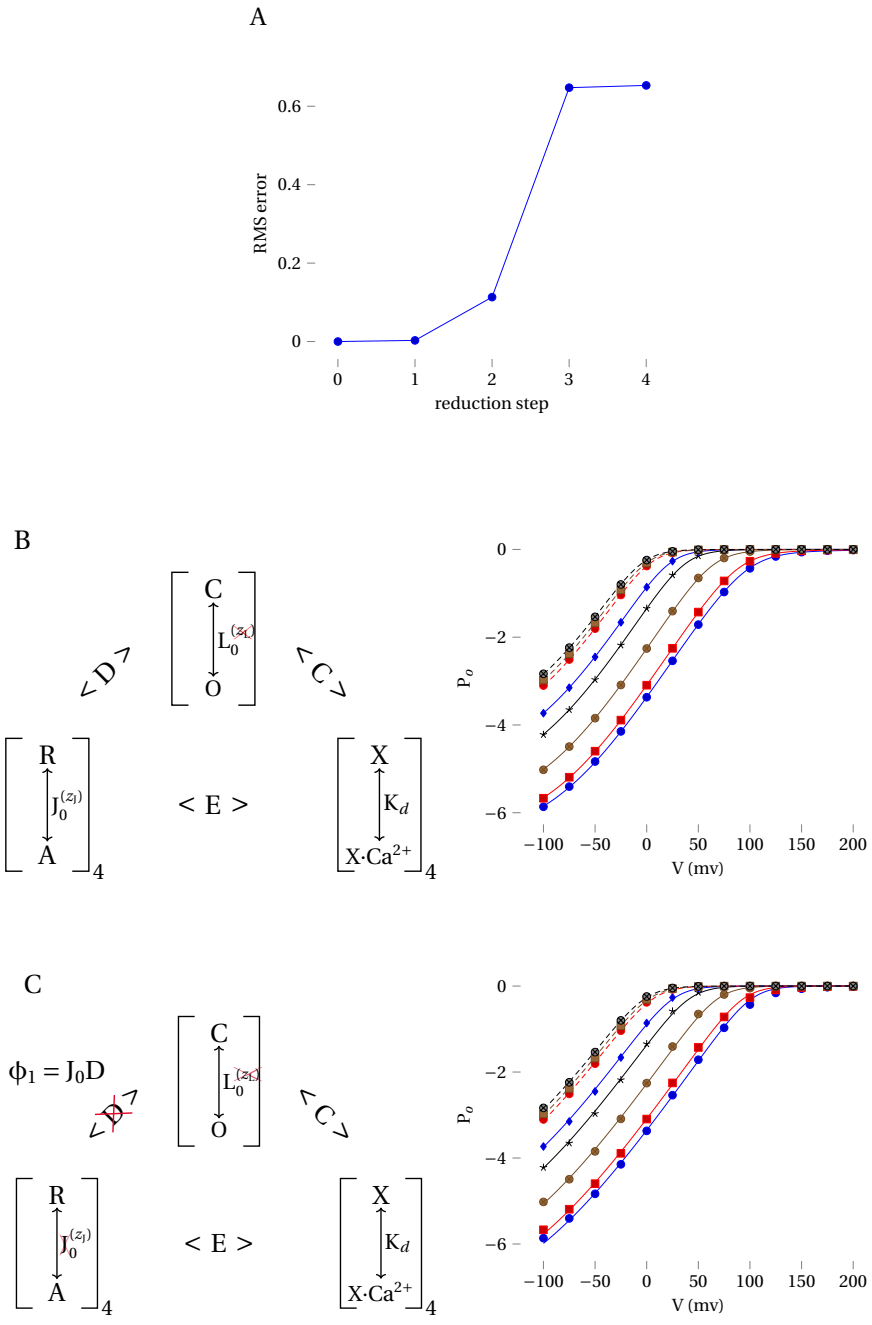
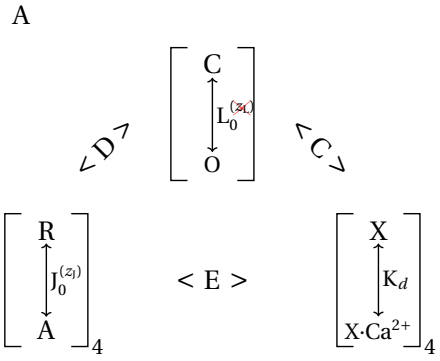
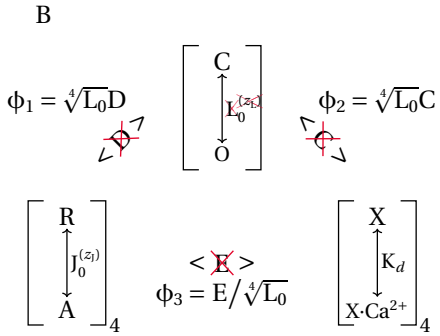


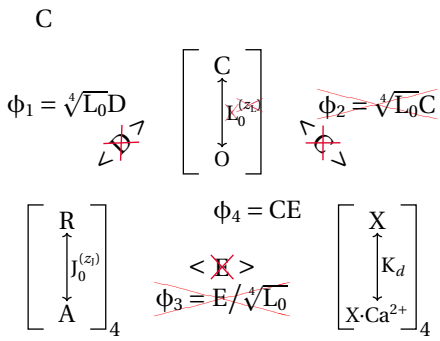
Figure 6: Overview of model reduction for log(P_o). (A) Error for four reduction steps. The first step produces essentially zero change in the fit, the second step produces a small increase in error. (B) Model resulting from one reduction (left) and fit to synthetic data (right, black lines through data points, data legend as in Fig 1). The model fits the data very well. (C). Model resulting from two reductions (left) and fit to synthetic data (right). The model fits data well, with small discrepancies for data at -50 to 50 mV and 22-70 μM Ca^{2+}.



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 7: Model reduction results in identifiable parameters. Reduced models for the P_o assay are presented in the left column (identical with Figure), and the error in their parameters (95% confidence interval) are presented at right. The five parameter model produced by three model reductions (third row, left) has fully identifiable parameters (error <10%, third row, right).

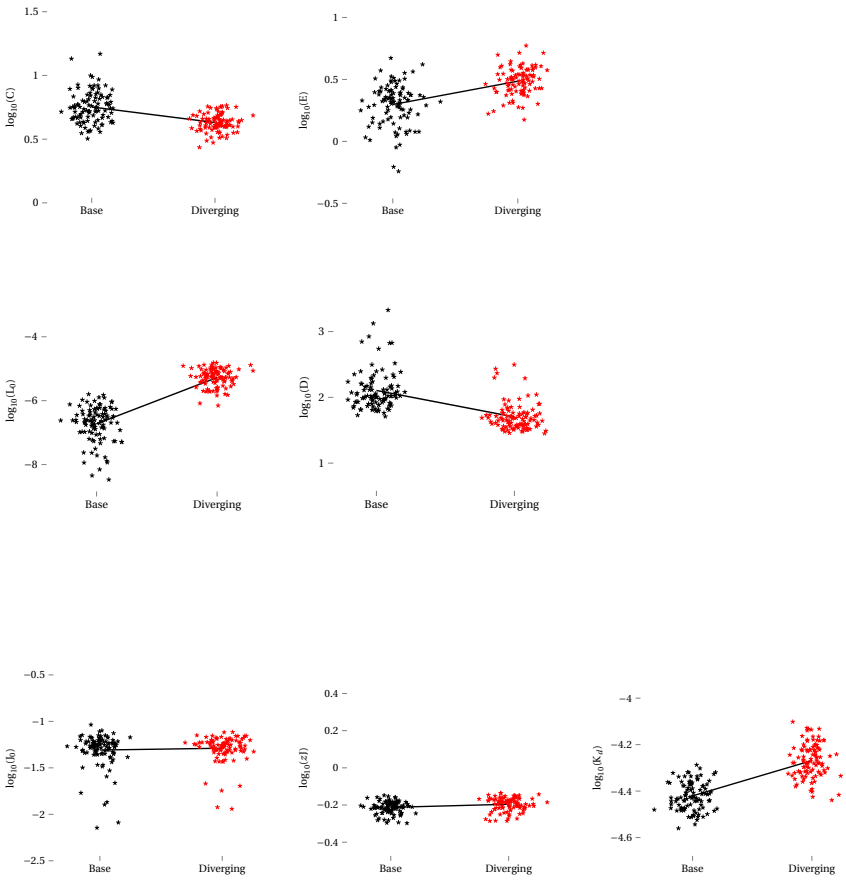


Figure 8: Identifying compensating parameters under noise