REGULARIZATION

$$V_N'' \sim \left(\sum \varphi \varphi^T \right)$$ CAN BE ILL CONDITIONED

$$W_N(\theta, Z^N) = \underbrace{\frac{1}{N} \sum_{t=1}^{N} \ell(\varepsilon_F(t,\theta))}_{V_N(\theta, Z^N)} + \delta |\theta - \theta_0|^2 \qquad \left(\underset{\text{GENERALLY}}{\theta_0 = \phi} \right)$$

$(V_N'' + \delta I)$ POSSIBLY BETTER CONDITIONED

FOR SHORT DATA → PARAMETERS OF SMALLEST INFLUENCE PULLED TO θ_0 (TO ϕ)

→ CONTROLLING THE EFFECTIVE NUMBER OF PARAMETERS

ASSUMPTION OF PROBABILISTIC FRAMEWORK: $\boxed{\text{DATA COLLECTION CAN BE REPEATED } \infty \text{ MANY TIMES UNDER "SIMILAR" CONDITIONS (STATIONARITY)}}$

BAYESIAN VIEW – MAX POSTERIOR ESTIMATE

θ RANDOM, PRIOR PDF — $g_\theta(z) = P(\theta = z)$

OBSERVATIONS RANDOM — $f_y(\theta, y^N)$

$$P(\theta | y^N) = \frac{P(y^N|\theta)\, P(\theta)}{P(y^N)} \sim f_y(\theta, y^N) \cdot g_\theta(\theta)$$

$$\hat{\theta}_{MAP}(y^N) = \arg\max_\theta \left[f_y(\theta, y^N)\, g_\theta(\theta) \right]$$

MAP → OPTIMAL REGULARIZATION $\sim \alpha / \beta$

PRIOR VARIANCE OF θ σ^2_{PRIOR}

~~VARIANCE OF~~ DISTURBANCES σ^2_{NOISE}