

対数線形モデルにおける最適な平坦 e-部分多様体

周徳潤, 杉山 磨人
(国立情報学研究所、総合研究大学院大学)

どんな研究？

- 入力分布の近似精度は、選択する e-平坦部分多様体によって異なる。
- 最も良い近似を与える e-平坦部分多様体を見つけたいと考える。

何がわかる？

- 近似精度が高い e-平坦部分多様体を探索する手法を提案する。
- 最も近似精度が高い e-平坦部分多様体は、0 制約付きであることが分かる。

背景・目的

離散分布の対数線形モデル化では、自然パラメータの一部にゼロの制約を課すことでモデル選択を行う。この手法は情報幾何において確立された概念であり、古典的なボルツマンマシンから最近のテンソルの多体近似まで幅広く用いられてきた。しかし、理論上はゼロ以外の定数値も制約に用いることが可能であり、これにより射影される部分多様体が変化する。本研究では、そのような制約値の漸近的挙動を情報幾何の視点から解析する。

研究内容（方法・結果・結論）

方法

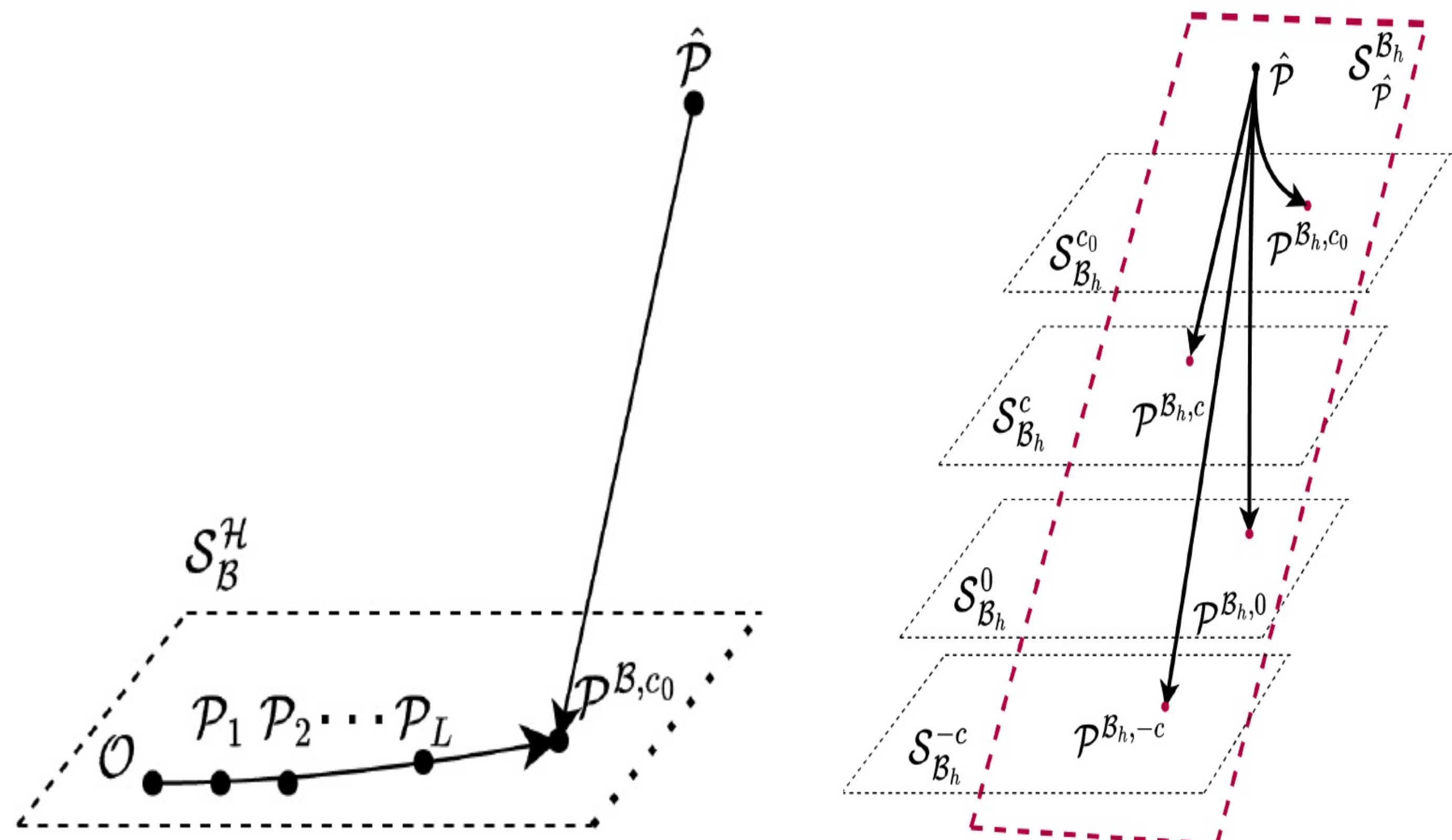
- 現在のパラメータで分布 $\mathcal{P}$ を計算
- η_k を更新
- 情報幾何に基づく行列 $\mathbf{G}$, $\tilde{\mathbf{G}}$ を構築
- 勾配 $\Delta\eta$ を計算
- $\tilde{\mathbf{G}}^{-1}$ を使ってパラメータを更新

凸最適化！



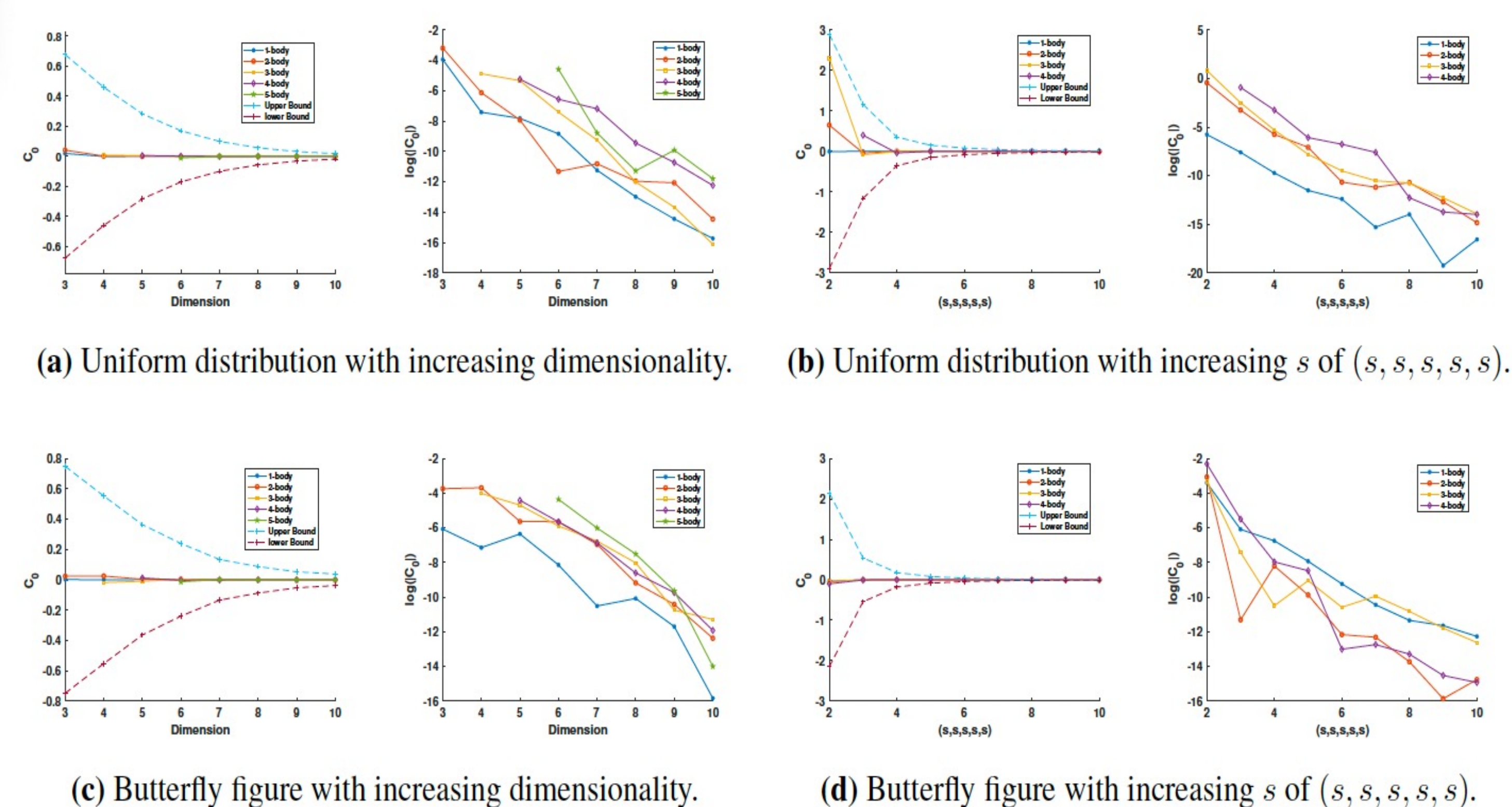
$$\mathcal{S}_B^H = \{ \mathcal{R} \in \mathcal{S} \mid \theta_\alpha = \theta_\beta \text{ for all } \alpha, \beta \in \Omega_d^+ \setminus \mathcal{B} \},$$

$$\mathcal{S}_{B_h}^c = \{ Q \in \mathcal{S} \mid \theta_v = c \text{ for all } v \in \Omega_d^+ \setminus \mathcal{B}_h \}$$



結果

漸近的挙動



理論解析

入力分布のサイズが大きくなるにつれて、最適な e-平坦部分多様体は 0 制約を持つ部分多様体に収束する。

$$\frac{2^d (d+1) I_j \log(\tau) b^2}{(I_j - m_j)(I_j - m_j + 1) \prod_{h=1, h \neq j}^d (1 + I_h) (a^2)},$$