pairwise model.

We will introduce algorithms that minimize the approximate objective $L(\delta)$ using local updates. Each iteration of the algorithms repeatedly finds a maximizing assignment for the subproblems individually, using these to update the dual variables that glue the subproblems together. We describe two classes of algorithms, one based on a subgradient method (see Section 1.4) and another based on block coordinate descent (see Section 1.5). These dual algorithms are simple and widely applicable to combinatorial problems in machine learning such as finding MAP assignments of graphical models.

1.3.1 Derivation of Dual

In what follows we show how the dual optimization in Eq. 1.2 is derived from the original MAP problem in Eq. 1.1. We first slightly reformulate the problem by duplicating the $x_i$ variables, once for each factor, and then enforce that these are equal. Let $x_i^f$ denote the copy of $x_i$ used by factor $f$. Also, denote by $x^f = \{x_i^f\}_{i \in f}$ the set of variables used by factor $f$, and by $x^F = \{x_f\}_{f \in F}$ the set of all variable copies. This is illustrated graphically in Fig. 1.3. Then, our reformulated – but equivalent – optimization problem

\begin{align*}
\min_{\delta_k} & \sum_{x_{ij}} \delta_k(x_{ij}) \\
\text{subject to} & \quad \sum_{x_{ij} \in f} \delta_f(x_{ij}) = \sum_{x_{ij} \in h} \delta_h(x_{ij}) = \sum_{x_{ij} \in k} \delta_k(x_{ij}) = 0, \\
& \quad \delta_{f1}(x_1) - \delta_{f2}(x_2) = \theta_f(x_1, x_2), \\
& \quad \delta_{h2}(x_2) - \delta_{h4}(x_4) = \theta_h(x_2, x_4), \\
& \quad \delta_{g1}(x_1) - \delta_{g3}(x_3) = \theta_g(x_1, x_3).
\end{align*}