This document is supplementary material for the paper entitled "Learning Scale-free Networks by Dynamic Node Specific Prior" accepted by ICML 2015. In this document, when we refer to one equation, theorem or conclusion in the main text, we would explicitly mention it. Otherwise, we refer to an equation, theorem or conclusion in this document.

1 PROOF OF EQUATION (8) IN MAIN TEXT

Equation (8) in the main text is equivalent to the claim that "The degree prior (7) in the main text favors a $p$ variable graph satisfying the given degree distribution $\{d_1, d_2, ..., d_p\}$”. Now we prove Equation 8 in the main text.

Proof. Assume the degree distribution of $E^{(2)}$ in Equation 8 is $\{d'_1, d'_2, ..., d'_p\}$ and inconsistent with the degree distribution $\{d_1, d_2, ..., d_p\}$. Without loss of generality, assume that $d'_1 > d_1$ and $d'_2 < d_2$. We can construct a graph $E'$ by moving one edge of variable 1 to variable 2. Actually we have

$$\sum_{u=1}^{p} \frac{H \circ E'_{u}}{d_u} - \sum_{u=1}^{p} \frac{H \circ E^{(2)}_{u}}{d_u} = \sum_{u=1}^{2} \frac{H \circ E'_{u}}{d_u} - \sum_{u=1}^{2} \frac{H \circ E^{(2)}_{u}}{d_u} = \frac{H_{d'_1}}{H_{d_1}} + \frac{H_{d'_2+1}}{H_{d_2}} < 0 \tag{1.1}$$

We can repeat such a construction process as long as $E'$ violates the degree distribution $\{d_1, d_2, ..., d_p\}$. Thus we have proved that Equation (8) in the main text holds. In other words, degree prior (7) in the main text favors graphs following the given degree distribution. \qed
2 PROOF OF THEOREM 2

Proof. Let \( Y^{t+1} \) denote the output of Algorithm 1. It satisfies \( [Y^{t+1}, H, \delta^t] = [Y^t, H, \delta^t] = [Y^{t+1}, h] \). So

\[
\arg\min_Y \frac{1}{2} \| Y - A \|_F^2 + \lambda \sum_{i=1}^{P} g(v) Y_{[i,Y^{t+1},h]} \circ H
\]  

(2.1)

is equivalent to

\[
\arg\min_Y \frac{1}{2} \| Y - A \|_F^2 + \lambda \sum_{i=1}^{P} g(v) Y_{[i,Y^{t+1},H,\delta^t]} \circ H
\]

(2.2)

Since \( [Y^{t+1}, H, \delta^t] = [Y^t, H, \delta^t] \), both \( \delta \) and the objective function would not change, which means \( Y^{t+1} \), the output of Algorithm 1 in the main text, is also the solution of (2.2) and (2.1). Thus Theorem 2 in the main text is proved.

\[\square\]

3 PROOF OF EQUATION 21 IN MAIN TEXT

In this section, we prove Eq. (21) in the main text. We also list necessary definitions and lemmas for the next section, where we prove the correctness of our algorithm.

Actually, each subproblem in Eq. (20) in the main text can be written as

\[
\frac{1}{2} \| Y_i - M_i \|_F^2 + \sum_{k=1}^{p-1} |Y_i,(k)| g(k)
\]  

(3.1)

Here, \( \{g(1), g(2), ..., g(p-1)\} \) is a sequence of positives. Given a feasible solution \( Y_i \), we can assume \( Y_{i,i} = M_{i,i} \), otherwise \( Y_i \) cannot be the optimal solution. When we sort the elements in \( Y_i \) and \( M_i \), we do not consider their \( i \)th elements \( Y_{i,i} \) and \( M_{i,i} \). That is, we sort all the elements in \( Y_i \) and \( M_i \) excluding \( Y_{i,i} \) and \( M_{i,i} \).

\[\text{Figure 3.1: Rank solution. The one-one correspondence is defined.}\]

As shown in Figure 3.1, we may sort all the elements of \( Y_i \) by their absolute values and obtain \( |Y_{i,(1)}| \geq |Y_{i,(2)}| \geq ... \geq |Y_{i,(p-1)}| \). Obviously, there is one-one correspondence between \( Y_{i,*} \) and \( Y_{i,(*)} \). If \( Y_{i,j} \) corresponds to \( Y_{i,(k)} \), then we denote the correspondence as \( \text{Map}(Y_{i,(k)}) = j \).

Assume that \( |Y_{i,(1)}| = |Y_{i,(2)}| = ... = |Y_{i,(t+1)}| \), we denote \( \{Y_{i,(1)}, Y_{i,(2)}, ..., Y_{i,(t)}\} \) as \( C|Y_i,(1)\). Similarly, we can construct \( C|Y_i,(2)\), \( C|Y_i,(3)\) and so on. Assume \( \bigcup_{k=1}^{p-1} Y_{i,(k)} \) can be clustered into
Given a feasible solution \( Y_i \) of (3.1), \( Y_i \) is a stable solution of (3.1) if and only if for each \( 1 \leq j \leq T \),

\[
y(j) = \max \{0, \frac{1}{|C(j)|} \sum_{s \in C(j)} (|M_i, Map(Y_{i,0})| - g(s))\}
\]  

(3.2)

Given Definition (1), we have the following two lemmas.

**Lemma 1.** The optimal solution \( Y_i^* \) of (3.1) is a stable solution.

**Proof.** First we rewrite the objective function (3.1) as follows.

\[
\frac{1}{2} \sum_{j=1}^{p} (|Y_{i,j}| - |M_{i,j}|)^2 + \sum_{k=1}^{p-1} |Y_{i,(k)}| g(k)
\]  

(3.3)

Given a \( j \) such that \( 1 \leq j \leq T(Y_i) \), if \( y(j) \neq 0 \), we can calculate the gradient over the absolute value of elements in \( C(k) \), and set it to 0, that is,

\[
\sum_{s \in C(j)} (|Y_{i,(s)}| - |M_{i, Map(Y_{i,0})}| + g(s)) = 0
\]  

(3.4)

Solving the above equation yields the following equation,

\[
y(j) = \frac{1}{|C(j)|} \sum_{s \in C(j)} (|M_{i, Map(Y_{i,0})}| - g(s))
\]  

(3.5)

If \( y(k) = 0 \), by calculating sub-gradient, it is easy to show that \( \sum_{s \in C(j)} (|M_{i, Map(Y_{i,0})}| - g(s)) \leq 0 \). Thus Lemma 1 is proved.

\[ \square \]

**Lemma 2.** Solving (3.3) is equivalent to the following optimization problem,

\[
\max \sum_{k=1}^{p-1} \hat{Y}_{i,(k)}
\]  

(3.6)

subject to the condition that \( \hat{Y}_i \) is a stable solution of (3.1) and (3.3).
Proof. According to Lemma 1, the optimal solution is stable solution, thus we just need to consider the set of all stable solutions when solving (3.1). Given a stable solution $\hat{Y}_i$, Substitute $\hat{Y}_i$ to (3.1) and exclude $Y_i$, we have following form,

$$
\frac{1}{2} ||\hat{Y}_i - M_i||_F^2 + \sum_{k=1}^{p-1} |\hat{Y}_{i,(k)}| g(k)
$$

$$= \frac{1}{2} \sum_{k=1}^{p-1} (|\hat{Y}_{i,(k)}| - |M_{i,(j)}|)^2 + \sum_{k=1}^{p-1} |\hat{Y}_{i,(k)}| g(k)
$$

$$= \frac{1}{2} \sum_{k=1}^{p-1} |\hat{Y}_{i,(k)}|^2 + \frac{p-1}{2} \sum_{k=1}^{p-1} |M_{i,(k)}|^2 - \sum_{k=1}^{p-1} |\hat{Y}_{i,(k)}| (|M_{i,(k)}| - g(k))
$$

$$= \frac{1}{2} \sum_{k=1}^{p-1} |\hat{Y}_{i,(k)}|^2 + \frac{1}{2} \sum_{k=1}^{p-1} |M_{i,(k)}|^2 - \sum_{j=1}^{\hat{p}} |\hat{C}(j)| \hat{y}(j)^2
$$

$$= \frac{1}{2} \sum_{k=1}^{p-1} |\hat{Y}_{i,(k)}|^2 + \frac{1}{2} \sum_{k=1}^{p-1} |M_{i,(k)}|^2 - \sum_{k=1}^{p-1} |\hat{Y}_{i,(k)}|^2
$$

$$= \frac{1}{2} \sum_{s=1}^{p-1} M_{i,(s)}^2 - \sum_{k=1}^{p-1} |\hat{Y}_{i,(k)}|^2
$$

(3.7)

As $\sum_{s=1}^{p-1} M_{i,(s)}^2$ is constant, so it is reasonable to conclude that solving (3.1) is equivalent to maximize $\sum_{k=1}^{p-1} |\hat{Y}_{i,(k)}|^2$ in the space of all stable solutions. Thus we conclude that Lemma 2 is correct.

Based on Lemma 2, we can define the partial order relationship among stable solutions.

Definition 2. Given two stable solutions $Y_i^{(1)}$ and $Y_i^{(2)}$ of (3.1), we say $Y_i^{(1)} \succ Y_i^{(2)}$ if and only if the following inequality holds

$$
\sum_{j=1}^{p-1} Y_i^{(1)}(j)^2 > \sum_{j=1}^{p-1} Y_i^{(2)}(j)^2
$$

(3.8)

Let $Y_i^*$ be the optimal solution. Then there is no feasible solution $Y_i$ such that $Y_i > Y_i^*$. Before proving Equation 21 in the main text, we show the following lemma.

Lemma 3. Let $\{x_1, x_2, ..., x_n\}$ be a non-negative non-decreasing sequence and $\{\delta_1, \delta_2, ..., \delta_n\}$ be a set of non-negative numbers satisfying $\sum_{i=1}^{s} \delta_i \geq \sum_{i=s+1}^{n} \delta_i > 0$. Then the following inequality holds,

$$
\sum_{i=1}^{s} (x_i + \delta_i)^2 + \sum_{i=s+1}^{n} (x_i - \delta_i)^2 \geq \sum_{i=1}^{n} x_i^2
$$

(3.9)
Lemma 5.

To prove

\begin{align*}
\sum_{i=1}^{s} (x_i^t + \delta_i)^2 + \sum_{i=s+1}^{n} (x_i - \delta_i)^2 &= \sum_{i=1}^{s} x_i^2 + \sum_{i=1}^{s} \delta_i^2 + 2[\sum_{i=s+1}^{s} x_i \delta_i - \sum_{i=s+1}^{n} x_i \delta_i] \\
&\geq \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} \delta_i^2 + 2[\sum_{i=s+1}^{s} x_i \delta_i - \sum_{i=s+1}^{n} \delta_i] \\
&= \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} \delta_i^2 + 2[(x_s - x_{s+1})] \sum_{i=s+1}^{n} \delta_i \\
&\geq \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} \delta_i^2 \\
&\geq \sum_{i=1}^{n} x_i^2
\end{align*}

Note that strict inequality holds as long as at least two elements in \( \{x_1, x_2, ..., x_n\} \) are different.

\[ \square \]

Actually, proving Equation 21 is equivalent to proving Lemma 4.

Lemma 4. Given the optimal solution \( Y_i^* \), for every \( j \) such that \( 1 \leq j \leq T^* \), the following two sets are equal,

\[ \{M_i, \text{Map}(Y, \Sigma_{x=1}^{n} |C(x)| + 1), M_i, \text{Map}(Y, \Sigma_{x=1}^{n} |C(x)| + 2), ..., M_i, \text{Map}(Y, \Sigma_{x=1}^{n} |C(x)| + 1)\} \]

\[ \{M_i, \Sigma_{x=1}^{n} |C(x)| + 1), M_i, \Sigma_{x=1}^{n} |C(x)| + 2), ..., M_i, \Sigma_{x=1}^{n} |C(x)|\} \]

Further, to prove Lemma 4, we just need to prove the following Lemma 5, which is much easier to prove.

Lemma 5. Let \( Y_i^{(0)} \) be a stable solution. If there exists a number \( q^{(0)} \) (\( 1 \leq q^{(0)} \leq T^{(0)} \)) and two numbers \( s_1, s_2 \) where \( \sum_{x=1}^{q^{(0)}} |C(x)| < s_1 \leq \sum_{x=1}^{q^{(0)}} |C(x)| \) and \( \sum_{x=1}^{q^{(0)}} |C(x)| < s_2 \leq \sum_{x=1}^{q^{(0)+1}} |C(x)| \), such that \( M_i, \text{Map}^{(q^{(0)})}(Y_i^{(0)}) < M_i, \text{Map}^{(q^{(0)+1})}(Y_i^{(0)}) \), then there is another stable solution \( Y_i' \), such that \( Y_i' > Y_i \). Here we use \( \text{Map}^{(t)} \) to denote the mapping of \( Y_i^{(t)} \).

Proof. To prove Lemma 5, we just need to show that we can construct a \( Y_i' \) described in Lemma 5.
We first construct $Y_i^{(1)}$ as follows. We set $Map_i^{(1)}(Y_i(s_1)) = Map_i^{(0)}(Y_i(s_2))$ and $Map_i^{(1)}(Y_i(s_2)) = Map_i^{(0)}(Y_i(s_1))$ and then set each component of $Y_i^{(1)}$ as follows,

$$Y_i^{(1)} = Y_i^{(0)}$$

$$Y_i^{(1)}_{i(s)} = Y_i^{(0)}_{i(s)} \text{ if } 1 \leq s \leq \sum_{x=1}^{q_{i+1}^{(0)}} |C_i^{(0)}(x)| \text{ or } \sum_{x=1}^{q_{i+1}^{(0)}} |C_i^{(0)}(x)| < s \leq n - 1$$

$$Y_i^{(1)}_{i(s)} = \frac{1}{|C_i^{(0)}(q_{i+1}^{(0)})|} \sum_{j \in C_i^{(0)(q_{i+1}^{(0)})}} |M_i^{(1)}(Y_i(s))| - g(j) \text{ if } \sum_{x=1}^{q_{i+1}^{(0)}} |C_i(x)| < s \leq \sum_{x=1}^{q_{i+1}^{(0)}} |C_i(x)|$$

If $Y_i^{(1)}$ is consistent, i.e., satisfying Condition (3.10), then $Y_i^{(1)}$ is a stable solution. In this case we can simply set $Y_i = Y_i^{(1)}$.

$$y_i^{(1)}(1) > ... > y_i^{(1)}(q_{i+1}^{(0)} - 1) > y_i^{(1)}(q_{i+1}^{(0)}) > y_i^{(1)}(q_{i+1}^{(0)} + 1) > y_i^{(1)}(q_{i+1}^{(0)} + 2) > ... > y_i^{(1)}(T(y_i^{(1)}))$$ (3.10)

Otherwise, let $q^{(1)} = q_{i+1}^{(0)}$ and $l^{(1)} = q_{i+1}^{(0)} + 1$, and then continue the following process. Assume that we have constructed $Y_i^{(k)}$. We say that $Y_i^{(k)}$ is "left-consistent" if

$$y_i^{(k)}(q^{(k)}) < y_i^{(k)}(q^{(k)} - 1).$$ (3.11)

And similarly, we say that $Y_i^{(k)}$ is "right-consistent" if

$$y_i^{(k)}(q^{(k)} + 2) < y_i^{(k)}(q^{(k)} + 1).$$ (3.12)

Then we start to construct $Y_i^{(k+1)}$. If $Y_i^{(k)}$ is not "left-consistent", set $q^{(k+1)} = q^{(k)} - 1$, merge $C^{(k)}(q^{(k)})$ and $C^{(k)}(q^{(k)} - 1)$ to one group $C^{(k+1)}(q^{(k+1)})$ and set $y^{(k+1)}(q^{(k+1)})$ as follows

$$y_i^{(k+1)}(q^{(k+1)}) = \frac{|C^{(k)}(q^{(k)} - 1)|y_i^{(k)}(q^{(k)} - 1) + |C^{(k)}(q^{(k)})|y_i^{(k)}(q^{(k)})}{|C^{(k)}(q^{(k)} - 1)| + |C^{(k)}(q^{(k)})|}$$ (3.13)

Otherwise, we do not need merge the two groups, but just set $q^{(k+1)} = q^{(k)}$ and $y^{(k+1)}(q^{(k+1)}) = y_i^{(k)}(q^{(k)})$.

Also, if $Y_i^{(k)}$ is not "right-consistent", then merge $C^{(k)}(q^{(k)} + 1)$ and $C^{(k)}(q^{(k)} + 2)$ into one group $C^{(k+1)}(q^{(k+1)} + 1)$. Set $l^{(k+1)} = l^{(k)} + 1$, and set $y^{(k+1)}(q^{(k+1)} + 1)$ as follows

$$y_i^{(k+1)}(q^{(k+1)} + 1) = \frac{|C^{(k)}(q^{(k)} + 1)|y_i^{(k)}(q^{(k)} + 1) + |C^{(k)}(q^{(k)} + 2)|y_i^{(k)}(q^{(k)} + 2)}{|C^{(k)}(q^{(k)} + 1)| + |C^{(k)}(q^{(k)} + 2)|}$$ (3.14)

Otherwise, set $l^{(k+1)} = l^{(k)}$ and $y^{(k+1)}(q^{(k+1)} + 1) = y_i^{(k)}(q^{(k)} + 1)$. We then construct the rest part of $Y_i^{(k+1)}$ as follows,
Then we get a new solution \( Y_{i}^{(k+1)} \). If \( Y_{i}^{(k+1)} \) is consistent, then we denote \( Y_{i}' = Y_{i}^{(k+1)} \). Otherwise, we continue the process described above to construct \( Y_{i}^{(k+2)} \). Note that we would definitely stop as such kind of construction process will not go on forever. When the process stopped, we would definitely get a stable solution \( Y_{i}' = Y_{i}^{(t(i))} \). According Lemma 2, to prove that \( Y_{i}^{(t(i))} > Y_{i} \) we just need to prove the following inequality holds,

\[
\sum_{a=q(i)}^{l(i)} |C^{(0)}(a)|y^{(0)}(a) - y^{(t(i))}(q(i))|y^{(t(i))}(q(i))| + |C^{(t(i)}(q(i) + 1)|y^{(t(i))}(q(i) + 1)^2 \geq 0
\] (3.15)

Using Lemma 3 we can prove that \( \sum_{a=q(i)}^{l(i)} Y_{i,a}^2 < \sum_{a=q(i)}^{l(i)} Y_{i,a}^2 \), thus we have

\[
\sum_{a=q(i)}^{l(i)} |C^{(0)}(a)|y^{(0)}(a) - y^{(t(i))}(q(i))|y^{(t(i))}(q(i))| + |C^{(t(i)}(q(i) + 1)|y^{(t(i))}(q(i) + 1)^2
\] (3.16)

According to (3.16), Lemma 5 is proved. Thus Lemma 4 and Equation 21 in the main text are also proved.

4 Algorithm 2 in Main Text (Edge Ranking Updating)

4.1 Partition Problem

Based on the conclusion of last section, solving (3.1) can be reduced to the following problem (4.1)

\[
\min \left( \sum_{i=1}^{p-1} |Y_{v(i)} - M_{v(i)}|^2 + \sum_{i=1}^{p-1} g(i)|Y_{v(i)}| \right).
\] (4.1)

This can be understood as partitioning ordered sequences \( (M_{a(1)}, M_{a(2)}, \ldots, M_{a(p-1)}) \) into optimal blocks, as shown in Figure 4.1, such that the square sum of the final solution is maximized, as claimed in Lemma 2. In Figure 4, we use \( M(i) \) to denote \( M_{a(i)} \) for some \( a \).
4.2 Property of Optimal Solution

Given a number \( t \) (\( 1 \leq t < p \)), let \( Y^*_v: t \) denote the optimal solution of the following sub-problem

\[
SUB(t) = \min \left\{ \sum_{i=1}^{t} |Y_v(i) - M_v(i)|^2 + \sum_{i=1}^{t} g(i)|Y_v(i)| \right\}
\]

(4.2)

Actually, the cluster structure \{\( C\)\} obtained in proving Lemma 1 can be used to prove that the optimal partition of \((M_v(1), M_v(2), ..., M_v(t))\) corresponds to the optimal solution of \((M_v(1), M_v(2), ..., M_v(t))\). Now we show two properties of the cluster structure.

1. Denote \( C|_{Y^*_v: (1)} = \{Y_v(a), Y_v(a+1), ..., Y_v(b)\} \). For any \( a \leq s < b \), we have

\[
\frac{\sum_{i=a+1}^{b} |M_v(i) - g(i)|}{s-a+1} \leq \frac{\sum_{i=a+1}^{b} |M_v(i)|}{s-a+1}.
\]

2. For any \( b < s < p \), we have

\[
\frac{\sum_{i=a+1}^{b} |M_v(i)|}{s-a+1} > \frac{\sum_{i=a+1}^{b} |M_v(i) - g(i)|}{s-b}.
\]

Here we show the proof of Property 1 and Property 2 can be proved similarly.

**Proof.** We assume that Property 1 does not hold, that there exists a \( s \), such that

\[
\frac{\sum_{i=a+1}^{b} |M_v(i)|}{s-a+1} > \frac{\sum_{i=a+1}^{b} |M_v(i) - g(i)|}{s-b}.
\]

For simplicity, let \( Y^*: v: t = \{C(1), C(2), ..., C(T)\} \). We can construct \( Y^{(1)} \) as follows.

\[
Y^{(1)} = \{C(1), C(2), ..., C(k - 1), C_{left}(k), C_{right}(k), C(k + 1), ..., C(T)\}
\]

(4.3)

Each element of \( C_{left}(k) \) equals to \( \frac{\sum_{i=a+1}^{b} |M_v(i)|}{s-a+1} \) while each element of \( C_{right}(k) \) equals to \( \frac{\sum_{i=a+1}^{b} |M_v(i) - g(i)|}{s-a+1} \). If (4.3) is consistent (i.e., satisfying (3.10)), then the construction process ends. Otherwise, we start to construct \( Y^{(2)} \). The process is very similar to what we have done in proving Lemma 5. If \( Y^{(1)} \) is not left-consistent (see (3.11)), then we combine \( C(k - 1) \) and \( C_{left}(k) \) and if \( Y^{(1)} \) is not right-consistent (see (3.12)), then we combine \( C_{right}(k) \) and \( C(k + 1) \). If \( Y^{(2)} \) is not consistent, we will continue such a process until a consistent solution \( Y^{(m)} \) is obtained. By using the same technique in proving Lemma 5, we can show that \( Y^{(m)} > Y^*: v: t \), which contradicts to the assumption that \( Y^*: v: t \) is optimal. \( \square \)
4.3 Correctness of Theorem 3 in main text

Based on the two properties, we now show that Theorem 3 in the main text is correct.

Proof. For simplicity, we still denote \( Y^*_x \) as \( \{C(1), C(2), ..., C(T)\} \), and \( C(T+1) = \{|M_{x,t+1}| - g(t+1)\} \). Assume that \( k \) is the largest value we mentioned in Theorem 3 in the main text. For any \( 1 \leq x \leq k \), \( \{C(1), ..., C(x-1), C_x\} \) would be a consistent solution of \( SUB(t+1) \) (see definition 4.2), where \( C_x \) is a set with \( \sum_{s=x}^{T+1}|C(s)| \) elements whose absolute value all equal to \( y_x \)

\[
y_x = \max \left\{0, \frac{\sum_{i \in [T+1]} C(i) \left( |M_{x,t(i)}| - g(v) \right)}{\sum_{j=x}^{T+1} |C(s)|} \right\} \quad (4.4)
\]

Obviously, among these solutions, \( \{C(1), ..., C(k-1), C_k\} \) is the largest one based on Lemma 2. We now just need to prove that the optimal solution \( Y^*_y \) of \( SUB(t+1) \) is actually one of the \( k \) solutions \( \{C(1), ..., C(x-1), C_x\} \) for \( 1 \leq x \leq k \).

Assume \( Y^*_y = \{C(1), C(2), C(b-1), C'(b), C'(b+1), ..., C'(T')\} \), which means the \( Y^*_y \) share the first \( b-1 \) clusters, and \( C'(b) \neq C(b) \). We would like to prove \( b = T' \). If \( b \neq T' \), there are two cases

1. \(|C'(b)| > |C(b)|\). Denote \( C'(b) = \{Y_{b(m)}, Y_{b(m+1)}, ..., Y_{b(q)}\} \) while \( C(b) = \{Y_{b(m)}, Y_{b(m+1)}, ..., Y_{b(q)}\} \).

   According to Property 2, we have \( \sum_{i=m}^{q} |M_{b(i)}| - g(i) > \sum_{i=m}^{q} |M_{b(i)}| - g(i) \); however, according to Property 1, we have \( \sum_{i=m}^{q} |M_{b(i)}| - g(i) \leq \sum_{i=m}^{q} |M_{b(i)}| - g(i) \). Contradiction! Thus, the assumption \(|C'(b)| > |C(b)|\) does not hold.

2. \(|C'(b)| < |C(b)|\). We would lead to similar contradiction! Thus, the assumption \(|C'(b)| < |C(b)|\) does not hold.

Note that \( C'(b) \neq C(b) \) is equivalent to \(|C'(b)| \neq |C(b)|\). So the only possibility is the assumption that \( b \neq T' \) is wrong. Thus \( b \) should equal to \( T' \), and this is equivalent to say that \( Y^*_y \) is always among the \( k \) solutions.

Proved. \(\Box\)

4.4 \( O(p \log(p)) \) Algorithm

Theorem 3 in the main text defines the problem structure suitable for dynamic programming. We further notice the following lemma, which is not hard to prove, leads to the \( O(p \log(p)) \) algorithm in the main text.

Lemma 6. For simplicity, we denote \( Y^*_y = \{C(1), C(2), ..., C(T)\} \) and \( C(T+1) = \{|M_{y,t+1}| - g(t+1)\} \). Given \( \{C(1), ..., C(x-1), C_x\} \), where \( C_x \) is a set with \( \sum_{s=x}^{T+1}|C(s)| \) elements having the same value as \( y_x \)

\[
y_x = \max \left\{0, \frac{\sum_{i \in [T+1]} C(i) \left( |M_{x,t(i)}| - g(v) \right)}{\sum_{j=x}^{T+1} |C(s)|} \right\} \quad (4.5)
\]
If \( y(x - 1) < y_x \), then for any \( s \) where \( x \leq s \leq T + 1 \), \( \{C(1), \ldots, C(s - 1), C_s\} \) is not consistent as \( y(s - 1) < y_s \).

Similarly, if \( y(x - 1) > y_x \), then for any \( s \) where \( 1 \leq s \leq x \), \( \{C(1), \ldots, C(s - 1), C_s\} \) is consistent as \( y(s - 1) > y_s \).

**Lemma 6** allows us to check the largest \( k \) value in \( O(\log(p)) \), which is the time complexity to build solution of \( SUB(t + 1) \) from \( Y_{v;1:t}^* \). Thus, we can construct a \( Y_v^* = Y_{v;1:p-1}^* \) with time complexity \( O(p \log(p)) \).