Supplementary Material

Qingming Tang

May 18, 2015

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. □

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

$$\arg\min_{Y} \frac{1}{2}||Y - A||_F^2 + \lambda \sum_{i=1}^{p} g(v) Y_{[v, Y^{t+1}, h]} \circ H \tag{2.1}$$

is equivalent to

$$\arg\min_{Y} \frac{1}{2}||Y - A||_F^2 + \lambda \sum_{i=1}^{p} g(v) Y_{[v, Y^{t+1}, H, \delta^t]} \circ H \tag{2.2}$$

Since $[Y^{t+1}, H, \delta^t] = [Y^t, H, \delta^t]$, both δ 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.

□

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

$$\frac{1}{2}||Y_i - M_i||_F^2 + \sum_{k=1}^{p-1} |Y_{i,(k)}| g(k) \tag{3.1}$$

Here, $\{g(1), g(2), ..., g(p-1)\}$ is a sequence of positives. Given a feasible solution Y_i, we can assume $Y_{ii} = M_{ii}$, 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 $M_{i,i}$.

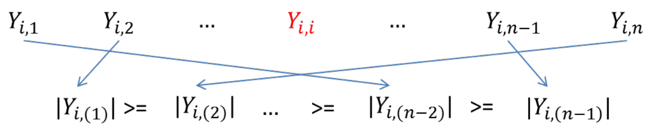

Figure 3.1: Rank solution. The one-one correspondence is defined.

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 $Map(Y_{i,(k)}) = j$.

Assume that $|Y_{i,(1)}| = |Y_{i,(2)}| = ... = |Y_{i,(t)}| > |Y_{i,(t+1)}|$, 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 $\cup_{k=1}^{p-1} Y_{i,(k)}$ can be clustered into

$T(Y_i)$ groups, that is $\{C|_{Y_i}(1), C|_{Y_i}(2), ..., C|_{Y_i}(T(Y_i))\}$. We use $|C|_{Y_i}(k)|$ to denote the number of elements in $C_{Y_i}(k)$, and for each element of $C|_{Y_i}(k)$, the absolute value should equal to $y|_{Y_i}(k)$. When context is clear, we use $C(k)$, $\hat{C}(k)$, $C'(k)$ and $C^{(l)}(k)$ for Y_i, $\hat{Y}_i$, Y_i' and $Y_i^{(l)}$, respectively, and use $y(k)$, T, $\hat{y}(k)$, $\hat{T}$, $y'(k)$, T', $y^{(l)}(k)$ and $T^{(l)}$ accordingly.

Definition 1. *Given a feasible solution Y_i of (3.1), Y_i is a stable solution of (3.1) if and only if for each $1 \le j \le T$,*

$$y(j) = max\{0, \frac{1}{|C(j)|}\sum_{s\in C(j)}[|M_{i,Map(Y_{i,(s)})}| - g(s)]\} \tag{3.2}$$

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.

$$\frac{1}{2}\sum_{j=1}^{p}(|Y_{i,j}| - |M_{i,j}|)^2 + \sum_{k=1}^{p-1}|Y_{i,(k)}|g(k) \tag{3.3}$$

Given a j such that $1 \le j \le T(Y_i)$, 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,

$$\sum_{s\in C(j)}\{|Y_{i,(s)}| - |M_{i,Map(Y_{i,(s)})}| + g(s)\} = 0 \tag{3.4}$$

Solving the above equation yields the following equation,

$$y(j) = \frac{1}{|C(j)|}\sum_{s\in C(j)}[|M_{i,Map(Y_{i,(s)})}| - g(s)] \tag{3.5}$$

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

□

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

$$Max_{\hat{Y}_i}\sum_{k=1}^{p-1}\hat{Y}_{i,(k)}^2 \tag{3.6}$$

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 $Y_i i$, we have following form,

$$\begin{aligned}
&\frac{1}{2}||\hat{Y}_i - M_i||_F^2 + \sum_{k=1}^{p-1} |\hat{Y}_{i,(k)}|g(k) \\
=&\frac{1}{2}\sum_{k=1}^{p-1} (|\hat{Y}_{i,(k)}| - |M_{i,(j)}|)^2 + \sum_{k=1}^{p-1} |\hat{Y}_{i,(k)}|g(k) \\
=&\frac{1}{2}\sum_{k=1}^{p-1} |\hat{Y}_{i,(k)}|^2 + \frac{1}{2}\sum_{k=1}^{p-1} |M_{i,(k)}|^2 - \sum_{k=1}^{p-1} |\hat{Y}_{i,(k)}|(|M_{i,(k)}| - g(k)) \\
=&\frac{1}{2}\sum_{k=1}^{p-1} |\hat{Y}_{i,(k)}|^2 + \frac{1}{2}\sum_{k=1}^{p-1} |M_{i,(k)}|^2 - \sum_{j=1}^{\hat{T}} |\hat{C}(j)|\hat{y}(j)^2 \\
=&\frac{1}{2}\sum_{k=1}^{p-1} |\hat{Y}_{i,(k)}|^2 + \frac{1}{2}\sum_{k=1}^{p-1} |M_{i,(k)}|^2 - \sum_{k=1}^{p-1} |\hat{Y}_{i,(k)}|^2 \\
=&\frac{1}{2}[\sum_{s=1}^{p-1} M_{i,(s)}^2 - \sum_{k=1}^{p-1} |\hat{Y}_{i,(k)}|^2]
\end{aligned} \tag{3.7}$$

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*

$$\sum_{j=1}^{p-1} Y_{i,(j)}^{(1)\ 2} > \sum_{j=1}^{p-1} Y_{i,(j)}^{(2)\ 2} \tag{3.8}$$

Let Y_i^* be the optimal solution. Then there is no feasible solution Y_i such that $Y_i \succ Y_i^*$. 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,*

$$\sum_{i=1}^{s} (x_i + \delta_i)^2 + \sum_{i=s+1}^{n} (x_i - \delta_i)^2 \geq \sum_{i=1}^{n} {x_i}^2 \tag{3.9}$$

Proof.

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

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

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,*

$$\{M_{i,Map(Y_{i,(\sum_{x=1}^{j-1}|C^*(x)|+1)})}, M_{i,Map(Y_{i,(\sum_{x=1}^{j-1}|C^*(x)|+2)})}, ..., M_{i,Map(Y_{i,(\sum_{x=1}^{j}|C^*(x)|)})}\}$$

$$\{M_{i,(\sum_{x=1}^{j-1}|C^*(x)|+1)}, M_{i,(\sum_{x=1}^{j-1}|C^*(x)|+2)}, ..., M_{i,(\sum_{x=1}^{j}|C^*(x)|)}\}$$

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 $\sum_{x=1}^{q^{(0)}-1}|C(x)| < s_1 \leq \sum_{x=1}^{q^{(0)}}|C(x)|$ and $\sum_{x=1}^{q^{(0)}}|C(x)| < s_2 \leq \sum_{x=1}^{q^{(0)}+1}|C(x)|$, such that $M_{i,Map^{(0)}(Y_{i,(s_1)})} < M_{i,Map^{(0)}(Y_{i,(s_2)})}$, then there is another stable solution Y_i', such that $Y_i' \succ Y_i$. Here we use $Map^{(t)}$ 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 $Map^{(1)}(Y_{i,(s_1)}) = Map^{(0)}(Y_{i,(s_2)})$ and $Map^{(1)}(Y_{i,(s_2)}) = Map^{(0)}(Y_{i,(s_1)})$ and then set each component of $Y_i^{(1)}$ as follows,

$$
\begin{aligned}
Y_{i,i}^{(1)} &= Y_{i,i}^{(0)} \\
Y_{i,(s)}^{(1)} &= Y_{i,(s)}^{(0)} \quad if \quad 1 \le s \le \sum_{x=1}^{q^{(0)}-1} |C^{(0)}(x)| \quad or \quad \sum_{x=1}^{q^{(0)}+1} |C^{(0)}(x)| < s \le n-1 \\
Y_{i,(s)}^{(1)} &= \frac{1}{|C^{(0)}(q^{(0)})|} \sum_{j \in C(q^{(0)})} [|M_{i,Map^{(1)}(Y_{i,(j)})}| - g(j)] \quad if \quad \sum_{x=1}^{q^{(0)}-1} |C(x)| < s \le \sum_{x=1}^{q^{(0)}} |C(x)| \\
Y_{i,(s)}^{(1)} &= \frac{1}{|C^{(0)}(q^{(0)}+1)|} \sum_{j \in C(q^{(0)}+1)} [|M_{i,Map^{(1)}(Y_{i,(j)})}| - g(j)] \quad if \quad \sum_{x=1}^{q^{(0)}} |C(x)| < s \le \sum_{x=1}^{q^{(0)}+1} |C(x)|
\end{aligned}
$$

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 $Y_i' = Y_i^{(1)}$.

$$y^{(1)}(1) > ... > y^{(1)}(q^{(0)}-1) > y^{(1)}(q^{(0)}) > y^{(1)}(q^{(0)}+1) > y^{(1)}(q^{(0)}+2) > ... > y^{(1)}(T(y^{(1)})) \quad (3.10)$$

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

$$y^{(k)}(q^{(k)}) < y^{(k)}(q^{(k)} - 1). \quad (3.11)$$

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

$$y^{(k)}(q^{(k)} + 2) < y^{(k)}(q^{(k)} + 1). \quad (3.12)$$

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

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

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

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

Otherwise, set $l^{(k+1)} = l^{(k)}$ and $y^{(k+1)}(q^{(k+1)} + 1) = y^{(k)}(q^{(k)} + 1)$. We then construct the rest part of $Y_i^{(k+1)}$ as follows,

$$
\begin{aligned}
Y_{i,i}^{(k+1)} &= Y_{i,i}^{(k)} \\
Y_{i,(s)}^{(k+1)} &= Y_{i,(s)}^{(k)} \quad if \quad 1 \le s \le \sum_{x=1}^{q^{(k+1)}-1} |C^{(k)}(x)| \\
Y_{i,(s)}^{(k+1)} &= Y_{i,(s)}^{(k)} \quad if \quad \sum_{x=1}^{q^{(k+1)}+1} |C^{(k+1)}(x)| < s < n
\end{aligned}
$$

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 $Y_i' = Y_i^{(t)}$. According **Lemma** 2, to prove that $Y_i^{(t)} \succ Y_i$ we just need to prove the following inequality holds,

$$
\sum_{a=q^{(t)}}^{l^{(t)}} |C^{(0)}(a)| y^{(0)}(a)^2 < |C^{(t)}(q^{(t)})| y^{(t)}(q^{(t)})^2 + |C^{(t)}(q^{(t)}+1)| y^{(t)}(q^{(t)}+1)^2 \tag{3.15}
$$

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

$$
\sum_{a=q^{(t)}}^{l^{(t)}} |C^{(0)}(a)| y^{(0)}(a)^2 = \sum_{a=q^{(t)}}^{l^{(t)}} Y_{i,a}^2 < \sum_{a=q^{(t)}}^{l^{(t)}} Y_{i,a}^{(t)2} = |C^{(t)}(q^{(t)})| y^{(t)}(q^{(t)})^2 + |C^{(t)}(q^{(t)}+1)| y^{(t)}(q^{(t)}+1)^2 \tag{3.16}
$$

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)

$$
\min\Big(\sum_{i=1}^{p-1} |Y_{v,(i)} - M_{v,(i)}|_F^2 + \sum_{i=1}^{p-1} g(i)|Y_{v,(i)}|\Big). \tag{4.1}
$$

This can be understood as partitioning ordered sequences ($M_{a,(1)}, M_{a,(2)}, ..., M_{a,(p-1)}$) 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.

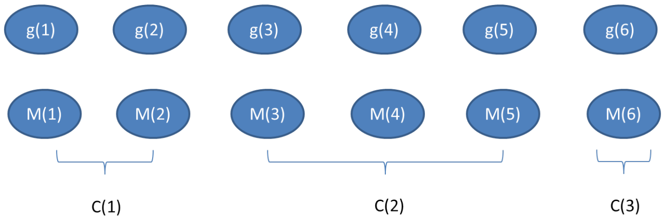

Figure 4.1: Partition of ordered sequences.

4.2 Property of optimal solution

Given a number t $(1 \le t < p)$, let $Y^*_{v,1:t}$ denote the optimal solution of the following subproblem

$$SUB(t) = \min\Big(\sum_{i=1}^{t} |Y_{v,(i)} - M_{v,(i)}|_F^2 + \sum_{i=1}^{t} g(i)|Y_{v,(i)}|\Big) \tag{4.2}$$

Actually, the cluster structure $\{C|_{Y^*_{v,1:t}}(1), C|_{Y^*_{v,1:t}}(2), ..., C|_{Y^*_{v,1:t}}(T(Y^*_{v,1:t}))\}$ of $Y^*_{v,1:t}$ corresponds to the optimal partition of $(M_{v,(1)},\ M_{v,(2)}, ..., M_{v,(t)})$. Now we show two properties of the cluster structure.

1 Denote $C|_{Y^*_{v,1:t}}(k) = \{Y_{v,(a)}, Y_{v,(a+1)}, ..., Y_{v,(b)}\}$. For any $a \le s < b$, we have $\frac{\sum_{i=a}^{s}|M_{v,(i)}|-g(i)}{s-a+1} \le \frac{\sum_{i=s+1}^{b}|M_{v,(i)}|-g(i)}{b-s}$.

2 For any $b < s < p$, we have $\frac{\sum_{i=a}^{b}|M_{v,(i)}|-g(i)}{b-a+1} > \frac{\sum_{i=b+1}^{s}|M_{v,(i)}|-g(i)}{s-b}$.

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 $\frac{\sum_{i=a}^{s}|M_{v,(i)}|-g(i)}{s-a+1} > \frac{\sum_{i=s+1}^{b}|M_{v,(i)}|-g(i)}{b-s}$.
For simplicity, let $Y^*_{v,1:t} = \{C(1), C(2), ..., C(T)\}$. We can construct $Y^{(1)}$ as follows.

$$Y^{(1)} = \{C(1), C(2), ..., C(k-1), C_{left}(k), C_{right}(k), C(k+1), ..., C(T)\} \tag{4.3}$$

Each element of $C_{left}(k)$ equals to $\frac{\sum_{i=a}^{s}|M_{v,(i)}|-g(i)}{s-a+1}$ while each element of $C_{right}(k)$ equals to $\frac{\sum_{i=s+1}^{b}|M_{v,(i)}|-g(i)}{b-s}$. 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 $Y^{(m)} \succ Y^*_{v,1:t}$, which contradicts to the assumption that $Y^*_{v,1:t}$ is optimal. □

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

$$y_x = \max\{0, \frac{\sum_{i \in \bigcup_{s=x}^{T+1} C(s)} \{|M_{v,(i)}| - g(v)\}}{\Sigma_{s=x}^{T+1}|C(s)|}\} \tag{4.4}$$

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 $Y^*_{v,1:t+1}$ of $SUB(t+1)$ is actually one of the k solutions $\{C(1), ..., C(x-1), C_x\}$ for $1 \le x \le k$.
Assume $Y^*_{v,1:t+1} = \{C(1), C(2), ..., C(b-1), C'(b), C'(b+1), ..., C'(T')\}$, which means the $Y^*_{v,1:t+1}$ and $Y^*_{v,1:t}$ share the first $b-1$ clusters, and $C'(b) \ne C(b)$. We would like to prove $b = T'$. If $b \ne T'$, there are two cases

1. $|C'(b)| > |C(b)|$. Denote $C'(b) = \{Y_{v,(m)}, Y_{v,(m+1)}, ..., Y_{v,(q)}\}$ while $C(b) = \{Y_{v,(m)}, Y_{v,(m+1)}, ..., Y_{v,(n)}\}$. According to **Property 2**, we have $\frac{\sum_{i=m}^{n}|M_{v,(i)}|-g(i)}{n-m+1} > \frac{\sum_{i=n+1}^{q}|M_{v,(i)}|-g(i)}{q-n}$; However, according to **Property 1**, we have $\frac{\sum_{i=m}^{n}|M_{v,(i)}|-g(i)}{n-m+1} \le \frac{\sum_{i=n+1}^{q}|M_{v,(i)}|-g(i)}{q-n}$. 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 $C'(b) \ne C(b)$ is equivalent to $|C'(b)| \ne |C(b)|$. So the only possibility is the assumption that $b \ne T'$ is wrong. Thus b should equal to T', and this is equivalent to say that $Y^*_{v,1:t+1}$ is always among the k solutions.
Proved. □

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* $Y^*_{v,1:t} = \{C(1), C(2), ..., C(T)\}$ *and* $C(T+1) = \{|M_{v,(t+1)}| - g(t+1)\}$. *Given* $\{C(1), ..., C(x-1), C_x\}$, *where* C_x *is a set with* $\Sigma_{s=x}^{T+1}|C(s)|$ *elements having the same value as* y_x

$$y_x = \max\{0, \frac{\sum_{i \in \bigcup_{s=x}^{T} C(s)} \{|M_{v,(i)}| - g(v)\}}{\Sigma_{s=x}^{T}|C(s)|}\} \tag{4.5}$$

If $y(x-1) < y_x$*, then for any* s *where* $x \leq s \leq T+1$, $\{C(1),...,C(s-1),C_s\}$ *is not consistent as* $y(s-1) < y_s$.
Similarly, if $y(x-1) > y_x$*, then for any* s *where* $1 \leq s \leq x$, $\{C(1),...,C(s-1),C_s\}$ *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 $O(p\log(p))$.